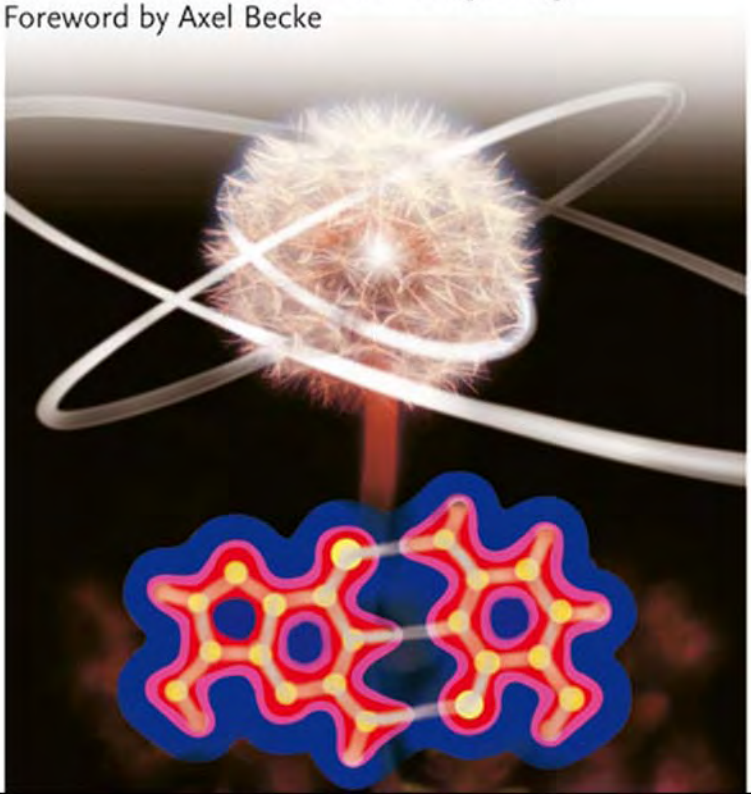
C. F. Matta and R. J. Boyd

WILEY-VCH

The Quantum Theory of Atoms in Molecules

From Solid State to DNA and Drug Design
Foreword by Axel Becke



The Quantum Theory of Atoms in Molecules

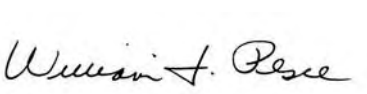
Edited by
Chérif F. Matta and
Russell J. Boyd

1807–2007 Knowledge for Generations

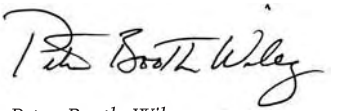
Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!



William J. Pesce
President and Chief Executive Officer



Peter Booth Wiley
Chairman of the Board

The Quantum Theory of Atoms in Molecules

From Solid State to DNA and Drug Design

Edited by
Chérif F. Matta and Russell J. Boyd

WILEY-VCH Verlag GmbH & Co. KGaA

The Editors

Prof. Chérif F. Matta
Department of Chemistry & Physics
Mount Saint Vincent University
Halifax, Nova Scotia, B3M 2J6
Canada
and
Department of Chemistry
Dalhousie University
Halifax, Nova Scotia, B3H 4J3
Canada

Prof. Russell J. Boyd
Department of Chemistry
Dalhousie University
Halifax, Nova Scotia, B3H 4J3
Canada

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
Die Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at ⟨http://dnb.d-nb.de⟩.

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Typesetting Asco Typesetters, Hong Kong
Printing betz-druck GmbH, Darmstadt
Binding Litges & Dopf GmbH, Heppenheim
Cover Design Adam-Design, Weinheim

ISBN 978-3-527-30748-7

This book is dedicated to

Professor Richard F.W. Bader

on the occasion of his 75th birthday,
for his invaluable contributions to the
physical theory of chemistry.

"... it is in virtue of the form that the matter is some one definite thing, and this is the substance of the thing. What Aristotle means seems to be plain common sense: a 'thing' must be bounded, and the boundary constitutes its form.... We should not naturally say that it is the form that confers substantiality, but that is because the atomic hypothesis is ingrained in our imagination. Each atom, however, if it is a 'thing', is so in virtue of its being delimited from other atoms, and so having, in some sense, a 'form'."

Bertrand Russell, *A History of Western Philosophy*;
Simon and Schuster, New York (1945).

Foreword

As a graduate student at McMaster University in the years 1975–81 I had the good fortune of knowing Richard Bader as a member of my supervisory committee. The departments of chemistry and physics shared facilities in the building that was then called the Senior Sciences Building and is now called the A.N. Bourns Building. I was a student in the department of physics, in the theoretical physics group, with an office not far from the offices of the Bader group.

Richard was the first theoretical chemist I ever met. In addition to the annual supervisory committee meetings, we chatted occasionally in his office. Looking back on those years, I wonder how differently my own career might have turned out had I not been inspired by those conversations. And inspired I was!

Anyone who knows Richard would agree that the single word describing him best is *passion*. Richard Bader is passionate about science. His conviction and enthusiasm for his own science, the theory of chemical bonding, are particularly infectious. Every session with him left me exhilarated and, at the same time, exhausted! His energy far outstripped mine.

What inspired me most was his passion for *ideas* and *concepts*. Our conversations were not about technical matters such as basis sets or levels of electron correlation. We talked about densities and probabilities in chemical systems, what makes a chemical bond, what is an *atom in a molecule*, why are atoms of a given element similar in different molecular environments? Fundamental questions! Fundamental ideas! My own efforts at the time were far off the beaten track – basis-set-free computational schemes, Thomas Fermi-like molecular models, the Hartree–Fock–Slater approximation. As a theoretical physics student with no expertise in theoretical chemistry, I relied heavily on Richard's feedback. Any other quantum chemist might have discouraged these unconventional projects, but not Richard. He was encouraging. He called the work important. He gave me confidence to pursue the unusual, ask the big questions, and, above all, seek out beautiful ideas.

Science is inherently beautiful. The most powerful ideas are also the simplest and the most beautiful. Early in the history of the density-functional theory of electronic structure, long before DFT was fashionable, Richard Bader followed his own path and developed some of the most beautiful ideas in chemistry. The boundaries of atoms in molecules are uniquely and rigorously defined by the

electron-density distribution. The topology of the density distribution also maps out all the bonds in a chemical system. Its Laplacian reveals the presence of localized electron pairs. Atoms in molecules, bonds, localized electron pairs – these are the most fundamental notions in chemistry. Richard Bader has taught us that the total electronic density defines them all.

It is a pleasure and an honor to help celebrate such a uniquely creative career. Thank you, Richard, for the inspiration and encouragement during my formative McMaster years and throughout my career. And thank you for the beautiful conceptual framework you gave to the theory of chemistry.

Dalhousie University, October 2006

Axel D. Becke
Killam Professor of
Computational Science, FRSC, FRS

Contents

Foreword *VII*

Preface *XIX*

List of Abbreviations Appearing in this Volume *XXVII*

List of Contributors *XXXIII*

1 An Introduction to the Quantum Theory of Atoms in Molecules *1*
Chérif F. Matta and Russell J. Boyd

1.1 Introduction *1*
1.2 The Topology of the Electron Density *1*
1.3 The Topology of the Electron Density Dictates the Form of Atoms in Molecules *5*
1.4 The Bond and Virial Paths, and the Molecular and Virial Graphs *8*
1.5 The Atomic Partitioning of Molecular Properties *9*
1.6 The Nodal Surface in the Laplacian as the Reactive Surface of a Molecule *10*
1.7 Bond Properties *10*
1.7.1 The Electron Density at the BCP (ρ_b) *11*
1.7.2 The Bonded Radius of an Atom (r_b), and the Bond Path Length *11*
1.7.3 The Laplacian of the Electron Density at the BCP ($\nabla^2\rho_b$) *11*
1.7.4 The Bond Ellipticity (ε) *12*
1.7.5 Energy Densities at the BCP *12*
1.7.6 Electron Delocalization between Bonded Atoms: A Direct Measure of Bond Order *13*
1.8 Atomic Properties *15*
1.8.1 Atomic Electron Population [$N(\Omega)$] and Charge [$q(\Omega)$] *16*
1.8.2 Atomic Volume [Vol.(Ω)] *16*
1.8.3 Kinetic Energy [$T(\Omega)$] *17*
1.8.4 Laplacian [$L(\Omega)$] *17*
1.8.5 Total Atomic Energy [$E_e(\Omega)$] *18*

1.8.6 Atomic Dipolar Polarization [$\mu(\Omega)$] *20*
1.8.7 Atomic Quadrupolar Polarization [$Q(\Omega)$] *24*
1.9 "Practical" Uses and Utility of QTAIM Bond and Atomic Properties *25*
1.9.1 The Use of QTAIM Bond Critical Point Properties *25*
1.9.2 The Use of QTAIM Atomic Properties *26*
1.10 Steps of a Typical QTAIM Calculation *27*
References *30*

Part I Advances in Theory *35*

2 The Lagrangian Approach to Chemistry *37*
Richard F. W. Bader

2.1 Introduction *37*
2.1.1 From Observation, to Physics, to QTAIM *37*
2.2 The Lagrangian Approach *38*
2.2.1 What is The Lagrangian Approach and What Does it Do? *38*
2.2.2 The Lagrangian and the Action Principle – A Return to the Beginnings *39*
2.2.3 Minimization of the Action *40*
2.2.4 Steps in Minimizing the Action *41*
2.3 The Action Principle in Quantum Mechanics *42*
2.3.1 Schrödinger's Appeal to the Action *42*
2.3.2 Schrödinger's Minimization *42*
2.3.2.1 Two Ways of Expressing the Kinetic Energy *43*
2.3.3 Obtaining an Atom from Schrödinger's Variation *44*
2.3.3.1 The Role of Laplacian in the Definition of an Atom *45*
2.3.4 Getting Chemistry from $\delta G(\psi,\nabla\psi;\Omega)$ *46*
2.4 From Schrödinger to Schwinger *48*
2.4.1 From Dirac to Feynman and Schwinger *48*
2.4.2 From Schwinger to an Atom in a Molecule *49*
2.5 Molecular Structure and Structural Stability *52*
2.5.1 Definition of Molecular Structure *52*
2.5.2 Prediction of Structural Stability *53*
2.6 Reflections and the Future *53*
2.6.1 Reflections *53*
2.6.2 The Future *55*
References *57*

3 Atomic Response Properties *61*
Todd A. Keith

3.1 Introduction *61*
3.2 Apparent Origin-dependence of Some Atomic Response Properties *62*
3.3 Bond Contributions to "Null" Molecular Properties *64*

3.4 Bond Contributions to Atomic Charges in Neutral Molecules *70*
3.5 Atomic Contributions to Electric Dipole Moments of Neutral Molecules *71*
3.6 Atomic Contributions to Electric Polarizabilities *73*
3.7 Atomic Contributions to Vibrational Infrared Absorption Intensities *78*
3.8 Atomic Nuclear Virial Energies *82*
3.9 Atomic Contributions to Induced Electronic Magnetic Dipole Moments *88*
3.10 Atomic Contributions to Magnetizabilities of Closed-Shell Molecules *90*
References *94*

4 QTAIM Analysis of Raman Scattering Intensities: Insights into the Relationship Between Molecular Structure and Electronic Charge Flow *95*
Kathleen M. Gough, Richard Dawes, Jason R. Dwyer, and Tammy L. Welshman

4.1 Introduction *95*
4.2 Background to the Problem *96*
4.2.1 Conceptual Approach to a Solution *97*
4.2.1.1 Experimental Measurement of Raman Scattering Intensities *97*
4.2.1.2 Theoretical Modeling of Raman Scattering Intensities: What We Did and Why *99*
4.3 Methodology *100*
4.3.1 Modeling α and $\partial\alpha/\partial r$ *101*
4.3.2 Recouping α From the Wavefunction, With QTAIM *102*
4.3.3 Recovering $\partial\alpha/\partial r$ From QTAIM *103*
4.4 Specific Examples of the Use of AIM2000 Software to Analyze Raman Intensities *103*
4.4.1 Modeling α in H_2 *104*
4.4.1.1 Modeling $\Delta\bar{\alpha}/\Delta r$ in H_2 *106*
4.4.2 Modeling α and $\Delta\bar{\alpha}/\Delta r$ in CH_4 *106*
4.4.3 Additional Exercises for the Interested Reader *108*
4.5 Patterns in α That Are Discovered Through QTAIM *109*
4.6 Patterns in $\partial\alpha/\partial r_{CH}$ That Apply Across Different Structures, Conformations, Molecular Types: What is Transferable? *111*
4.6.1 Patterns in $\Delta\bar{\alpha}/\Delta r_{CH}$ Revealed by QTAIM *111*
4.6.1.1 QTAIM Analysis of $\Delta\bar{\alpha}/\Delta r_{CH}$ in Small Alkanes *111*
4.6.1.2 What Did We Learn From QTAIM That Can be Transferred to the Other Molecules? *113*
4.7 What Can We Deduce From Simple Inspection of $\partial\bar{\alpha}/\partial r_{CH}$ and $\partial\bar{\alpha}/\partial r_{CC}$ From Gaussian? *114*
4.7.1 Variations in $\partial\bar{\alpha}/\partial r_{CH}$ Among the Alkanes *114*
4.7.2 $\Delta\bar{\alpha}/\Delta r_{CH}$ in Cycloalkanes, Bicycloalkanes, and Hedranes *116*
4.7.3 Patterns That Emerge in $\Delta\bar{\alpha}/\Delta r_{CC}$ of Alkanes *116*

4.7.4 Unsaturated Hydrocarbons and the Silanes: C–H, C=C, and Si–Si Derivatives *117*
4.8 Conclusion *118*
References *119*

5 Topological Atom–Atom Partitioning of Molecular Exchange Energy and its Multipolar Convergence *121*
Michel Rafat and Paul L. A. Popelier

5.1 Introduction *121*
5.2 Theoretical Background *123*
5.3 Details of Calculations *128*
5.4 Results and Discussion *130*
5.4.1 Convergence of the Exchange Energy *130*
5.4.2 Convergence of the Exchange Force *136*
5.4.3 Diagonalization of a Matrix of Exchange Moments *136*
5.5 Conclusion *139*
References *139*

6 The ELF Topological Analysis Contribution to Conceptual Chemistry and Phenomenological Models *141*
Bernard Silvi and Ronald J. Gillespie

6.1 Introduction *141*
6.2 Why ELF and What is ELF? *142*
6.3 Concepts from the ELF Topology *144*
6.3.1 The Synaptic Order *145*
6.3.2 The Localization Domains *145*
6.3.3 ELF Population Analysis *147*
6.4 VSEPR Electron Domains and the Volume of ELF Basins *149*
6.5 Examples of the Correspondence Between ELF Basins and the Domains of the VSEPR Model *153*
6.5.1 Octet Molecules *153*
6.5.1.1 Hydrides (CH_4, NH_3, H_2O) *153*
6.5.1.2 AX_4 (CH_4, CF_4, $SiCl_4$) *154*
6.5.1.3 AX_3E and AX_2E_2 (NCl_3, OCl_2) *154*
6.5.2 Hypervalent Molecules *155*
6.5.2.1 PCl_5 and SF_6 *155*
6.5.2.2 SF_4 and ClF_3 *155*
6.5.2.3 AX_7 and AX_6E Molecules *155*
6.5.3 Multiple Bonds *156*
6.5.3.1 C_2H_4 and C_2H_2 *156*
6.5.3.2 Si_2Me_4 and Si_2Me_2 *157*
6.6 Conclusions *158*
References *159*

Part II Solid State and Surfaces *163*

7 Solid State Applications of QTAIM and the Source Function – Molecular Crystals, Surfaces, Host–Guest Systems and Molecular Complexes *165*
Carlo Gatti

7.1 Introduction *165*
7.2 QTAIM Applied to Solids – the TOPOND Package *166*
7.2.1 QTAIM Applied to Experimental Densities: TOPXD and XD Packages *168*
7.3 QTAIM Applied to Molecular Crystals *170*
7.3.1 Urea *171*
7.3.1.1 Urea: Packing Effects *172*
7.4 QTAIM Applied to Surfaces *179*
7.4.1 Si(111)(1 × 1) Clean and Hydrogen-covered Surfaces *180*
7.4.2 Si(111)(2 × 1) Reconstructed Surface *184*
7.5 QTAIM Applied to Host–Guest Systems *186*
7.5.1 Type I Inorganic Clathrates $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba) *186*
7.5.2 Sodium Electrosodalite *190*
7.6 The Source Function: Theory *192*
7.6.1 The Source Function and Chemical Transferability *194*
7.6.2 Chemical Information from the Source Function: Long and Short-range Bonding Effects in Molecular Complexes *196*
7.6.3 The Source Function: Latest Developments *201*
References *202*

8 Topology and Properties of the Electron Density in Solids *207*
Víctor Luaña, Miguel A. Blanco, Aurora Costales, Paula Mori-Sánchez, and Angel Martín Pendás

8.1 Introduction *207*
8.2 The Electron Density Topology and the Atomic Basin Shape *209*
8.3 Crystalline Isostructural Families and Topological Polymorphism *213*
8.4 Topological Classification of Crystals *215*
8.5 Bond Properties – Continuity from the Molecular to the Crystalline Regime *217*
8.6 Basin Partition of the Thermodynamic Properties *219*
8.7 Obtaining the Electron Density of Crystals *222*
References *227*

9 Atoms in Molecules Theory for Exploring the Nature of the Active Sites on Surfaces *231*
Yosslen Aray, Jesus Rodríguez, and David Vega

9.1 Introduction *231*
9.2 Implementing the Determination of the Topological Properties of $\rho(r)$ from a Three-dimensional Grid *231*

9.3 An Application to Nanocatalyts – Exploring the Structure of the Hydrodesulfurization MoS_2 Catalysts *236*
9.3.1 Catalyst Models *237*
9.3.2 The Full $\rho(r)$ Topology of the MoS_2 Bulk *241*
9.3.3 The $\rho(r)$ Topology of the MoS_2 Edges *245*
References *254*

Part III Experimental Electron Densities and Biological Molecules *257*

10 Interpretation of Experimental Electron Densities by Combination of the QTAMC and DFT *259*
Vladimir G. Tsirelson

10.1 Introduction *259*
10.2 Specificity of the Experimental Electron Density *261*
10.3 Approximate Electronic Energy Densities *262*
10.3.1 Kinetic and Potential Energy Densities *262*
10.3.2 Exchange and Correlation Energy Densities *271*
10.4 The Integrated Energy Quantities *275*
10.5 Concluding Remarks *276*
References *278*

11 Topological Analysis of Proteins as Derived from Medium and High-resolution Electron Density: Applications to Electrostatic Properties *285*
Laurence Leherte, Benoît Guillot, Daniel P. Vercauteren, Virginie Pichon-Pesme, Christian Jelsch, Angélique Lagoutte, and Claude Lecomte

11.1 Introduction *285*
11.2 Methodology and Technical Details *287*
11.2.1 Ultra-high X-ray Resolution Approach *287*
11.2.2 Medium-resolution Approach *289*
11.2.2.1 Promolecular Electron Density Distribution Calculated from Structure Factors *289*
11.2.2.2 Promolecular Electron Density Distribution Calculated from Atoms *290*
11.2.3 A Test System – Human Aldose Reductase *291*
11.3 Topological Properties of Multipolar Electron Density Database *294*
11.4 Analysis of Local Maxima in Experimental and Promolecular Medium-resolution Electron Density Distributions *298*
11.4.1 Experimental and Promolecular Electron Density Distributions Calculated from Structure Factors *299*
11.4.2 Promolecular Electron Density Distributions Calculated from Atoms (PASA Model) *301*
11.5 Calculation of Electrostatic Properties from Atomic and Fragment Representations of Human Aldose Reductase *305*
11.5.1 Medium- and High-resolution Approaches of Electrostatic Potential Computations *307*

11.5.2 Electrostatic Potential Comparisons *309*
11.5.3 Electrostatic Interaction Energies *312*
11.6 Conclusions and Perspectives *312*
References *314*

12 Fragment Transferability Studied Theoretically and Experimentally with QTAIM – Implications for Electron Density and Invariom Modeling *317*
Peter Luger and Birger Dittrich

12.1 Introduction *317*
12.2 Experimental Electron-density Studies *318*
12.2.1 Experimental Requirements *318*
12.2.2 Recent Experimental Advances *319*
12.2.2.1 Synchrotron Radiation Compared with Laboratory Sources *319*
12.2.2.2 Data Collection at Ultra-low Temperatures (10–20 K) *321*
12.3 Studying Transferability with QTAIM – Atomic and Bond Topological Properties of Amino Acids and Oligopeptides *323*
12.4 Invariom Modeling *328*
12.4.1 Invariom Notation, Choice of Model Compounds, and Practical Considerations *330*
12.4.2 Support for Pseudoatom Fragments from QTAIM *331*
12.5 Applications of Aspherical Invariom Scattering Factors *334*
12.5.1 Molecular Geometry and Anisotropic Displacement Properties *334*
12.5.2 Using the Enhanced Multipole Model Anomalous Dispersion Signal *335*
12.5.3 Modeling the Electron Density of Oligopeptide and Protein Molecules *336*
12.6 Conclusion *338*
References *339*

Part IV Chemical Bonding and Reactivity *343*

13 Interactions Involving Metals – From "Chemical Categories" to QTAIM, and Backwards *345*
Piero Macchi and Angelo Sironi

13.1 Introduction *345*
13.2 The Electron Density in Isolated Metal Atoms – Hints of Anomalies *345*
13.3 Two-center Bonding *349*
13.3.1 The Dative Bond *350*
13.3.1.1 Metal Carbonyls *351*
13.3.1.2 Donor–Acceptor Interactions of Heavy Elements *352*
13.3.2 Direct Metal–Metal Bonding *352*
13.4 Three-center Bonding *356*
13.4.1 π-Complexes *357*
13.4.2 σ-Complexes *363*

13.4.2.1 Dihydrogen and Dihydride Coordination *364*
13.4.2.2 Agostic Interactions *364*
13.4.2.3 Hydride Bridges *367*
13.4.3 Carbonyl-supported Metal–Metal Interactions *370*
13.5 Concluding Remarks *371*
References *372*

14 Applications of the Quantum Theory of Atoms in Molecules in Organic Chemistry – Charge Distribution, Conformational Analysis and Molecular Interactions *375*
Jesús Hernández-Trujillo, Fernando Cortés-Guzmán, and Gabriel Cuevas

14.1 Introduction *375*
14.2 Electron Delocalization *375*
14.2.1 The Pair-density *375*
14.2.2 $^3J_{HH}$ Coupling Constants and Electron Delocalization *378*
14.3 Conformational Equilibria *380*
14.3.1 Rotational barriers *380*
14.3.1.1 Rotational Barrier of Ethane *380*
14.3.1.2 Rotational Barrier of 1,2-Disubstituted Ethanes *382*
14.3.2 Anomeric Effect on Heterocyclohexanes *386*
14.4 Aromatic Molecules *391*
14.4.1 Electronic Structure of Polybenzenoid Hydrocarbons *391*
14.5 Conclusions *395*
References *396*

15 Aromaticity Analysis by Means of the Quantum Theory of Atoms in Molecules *399*
Eduard Matito, Jordi Poater, and Miquel Solà

15.1 Introduction *399*
15.2 The Fermi Hole and the Delocalization Index *401*
15.3 Electron Delocalization in Aromatic Systems *403*
15.4 Aromaticity Electronic Criteria Based on QTAIM *404*
15.4.1 The *para*-Delocalization Index (PDI) *404*
15.4.2 The Aromatic Fluctuation Index (FLU) *406*
15.4.3 The π-Fluctuation Aromatic Index (FLU_π) *407*
15.5 Applications of QTAIM to Aromaticity Analysis *409*
15.5.1 Aromaticity of Buckybowls and Fullerenes *409*
15.5.2 Effect of Substituents on Aromaticity *412*
15.5.3 Assessment of Clar's Aromatic π-Sextet Rule *416*
15.5.4 Aromaticity Along the Diels–Alder Reaction. The Failure of Some Aromaticity Indexes *418*
15.6 Conclusions *419*
References *421*

16 Topological Properties of the Electron Distribution in Hydrogen-bonded Systems *425*
Ignasi Mata, Ibon Alkorta, Enrique Espinosa, Elies Molins, and José Elguero

16.1 Introduction *425*
16.2 Topological Properties of the Hydrogen Bond *426*
16.2.1 Topological Properties at the Bond Critical Point (BCP) *426*
16.2.2 Integrated Properties *429*
16.3 Energy Properties at the Bond Critical Point (BCP) *431*
16.4 Topological Properties and Interaction Energy *435*
16.5 Electron Localization Function, $\eta(r)$ *438*
16.6 Complete Interaction Range *440*
16.6.1 Dependence of Topological and Energy Properties on the Interaction Distance *440*
16.6.2 Perturbed Systems *448*
16.7 Concluding Remarks *450*
References *450*

17 Relationships between QTAIM and the Decomposition of the Interaction Energy – Comparison of Different Kinds of Hydrogen Bond *453*
Sławomir J. Grabowski

17.1 Introduction *453*
17.2 Diversity of Hydrogen-bonding Interactions *456*
17.3 The Decomposition of the Interaction Energy *459*
17.4 Relationships between the Topological and Energy Properties of Hydrogen Bonds *460*
17.5 Various Other Interactions Related to Hydrogen Bonds *464*
17.5.1 $H^+\cdots\pi$ Interactions *464*
17.5.2 Hydride Bonds *466*
17.6 Summary *467*
References *468*

Part V Application to Biological Sciences and Drug Design *471*

18 QTAIM in Drug Discovery and Protein Modeling *473*
Nagamani Sukumar and Curt M. Breneman

18.1 QSAR and Drug Discovery *473*
18.2 Electron Density as the Basic Variable *474*
18.3 Atom Typing Scheme and Generation of the Transferable Atom Equivalent (TAE) Library *476*
18.4 TAE Reconstruction and Descriptor Generation *478*
18.5 QTAIM-based Descriptors *480*
18.5.1 TAE Descriptors *482*
18.5.2 RECON Autocorrelation Descriptors *485*
18.5.3 PEST Shape–Property Hybrid Descriptors *485*

18.5.4 Electron Density-based Molecular Similarity Analysis *487*
18.6 Sample Applications *489*
18.6.1 QSAR/QSPR with TAE Descriptors *489*
18.6.2 Protein Modeling with TAE Descriptors *491*
18.7 Conclusions *492*
References *494*

19 Fleshing-out Pharmacophores with Volume Rendering of the Laplacian of the Charge Density and Hyperwall Visualization Technology *499*
Preston J. MacDougall and Christopher E. Henze

19.1 Introduction *499*
19.2 Computational and Visualization Methods *501*
19.2.1 Computational Details *501*
19.2.2 Volume Rendering of the Laplacian of the Charge Density *501*
19.2.3 The Hyperwall *505*
19.2.4 Hyper-interactive Molecular Visualization *505*
19.3 Subatomic Pharmacophore Insights *507*
19.3.1 Hydrogen-bonding Donor Sites *507*
19.3.2 Inner-valence Shell Charge Concentration (i-VSCC) Features in Transition-metal Atoms *509*
19.3.3 Misdirected Valence in the Ligand Sphere of Transition-metal Complexes *511*
19.4 Conclusion *513*
References *514*

Index *515*

Preface

"The manner in which the electron density is disposed in a molecule has not received the attention its importance would seem to merit. Unlike the energy of a molecular system which requires a knowledge of the second-order density matrix for its evaluation [a] many of the observable properties of a molecule are determined in whole or in part by the simple three-dimensional electron-density distribution. In fact, these properties provide a direct measure of a wide spectrum of different moments averaged directly over the density distribution. Thus the diamagnetic susceptibility, the dipole moment, the diamagnetic contribution to the nuclear screening constant, the electric field, and the electric field gradient (as obtained from nuclear quadrupole coupling constants) provide a measure of (aside from any angular dependencies) $\langle r_i^2 \rangle$, $\langle r_i \rangle$, $\langle r_i^{-1} \rangle$, $\langle r_i^{-2} \rangle$, *and* $\langle r_i^{-3} \rangle$, *respectively. The electric field at a nucleus due to the electron density distribution is of particular interest due to the theorem derived by Hellmann [b] and Feynman [c]. They have demonstrated that the force acting on a nucleus in a molecule is determined by the electric field at that nucleus due to the other nuclei and to the electron-density distribution."*

a P.-O. Löwdin, *Adv. Chem. Phys.* **2**, 207 (1959)
b J. Hellman, *Einführung in die Quantenchemie* (Detiche, Leipzig, Germany, 1937)
c R.P. Feynman, *Phys. Rev.* **56**, 340 (1939)

Richard F.W. Bader and Glenys A. Jones (1963) [a]

It has been sixteen years since the publication of Richard Bader's classic 1990 treatise *"Atoms in Molecules: A Quantum Theory"* [b]. The theory was founded on the recognition that the electron density plays a critical role in explaining and understanding the experimental observations of chemistry. Bader's work is among the earliest to draw attention to the importance of the electron density in chemistry, as the opening quotation, predating the discovery of the Hohenberg–Kohn

theorem, suggests. This 1963 paper includes an early example of molecular electron density contour plots (of the ammonia molecule).

Bader's fundamental work in the sixties on molecular electron density distributions (Table 1) laid the foundations for the theory which was developed in the seventies and eighties by his research group, which became known as the theory of atoms in molecules (AIM). In more recent literature this theory is often called the quantum theory of atoms in molecules (QTAIM) in recognition of its rigorous basis in quantum mechanics [2–6]. The theory relates the concepts of chemistry, for example chemical structure, chemical bonding, transferability of functional groups, and chemical reactivity, to the topology of the underlying electron-density distribution(s). QTAIM has, in effect, moved theoretical chemistry into real three-dimensional space [7]. In Bader's words:

> *"the charge* [electron] *density provides a description of the distribution of charge throughout real space and is the bridge between the concept of state functions in Hilbert space and the physical model of matter in real space."* [2]

By defining "proper open quantum systems" as special bounded regions within a closed (whole) system, followed by the identification of these regions as "atoms in molecules", the quantum theory of atoms in molecules brought quantum mechanics into applicability to an atom *within* a molecule. When a molecular property can be expressed in terms of a property *density*, the contribution of an atom to that molecular property can be obtained by integrating this density over the bounded volume of that atom in the molecule. In this way every atom in a molecule or crystal is characterized by a set of physical properties, each of which corresponds to a molecular property. These atomic properties, naturally, add up to those of the total molecular system and, for this reason, parallel and recover the properties of the atoms of experimental chemistry [8]. In this sense, the quantum theory of atoms in molecules is *the* quantum mechanics of atoms within molecules and crystals [9–11].

The virial theorem, which governs the relationship between the potential and kinetic energies of a molecule, occupies a prominent place in molecular quantum mechanics. This theorem has been generalized by Bader from its *global* statement (which applies to the molecule as a whole) to a *local* statement defined at every point in space [10]. In other words, the theorem has been re-written in its most general form which applies at every point of space in terms of scalar functions of space, i.e. densities. This very important generalization, known as the "local statement of the virial theorem", Eq. (54) in Ref. [10] and Eq. (10) in Chapter 1, is, perhaps, and to the best of our knowledge, the *only* known local relationship between the energy densities and the electron density that applies everywhere in space. More precisely, the local virial theorem relates the potential energy density and the kinetic energy density distributions locally to a function of the electron density, namely, its Laplacian [2, 10]. Bader also postulated [12], and later showed [2, 13–15] that the integrated form of this theorem, discovered before its local expression, translates into a virial theorem satisfied by each atom within a molecule

Table 1 Early Publications (nineteen-sixties) on molecular electron density distributions by Professor Richard F.W. Bader.

1. R.F.W. Bader and G.A. Jones, ''The Hellmann-Feynman Theorem and Chemical Binding'', *Canadian Journal of Chemistry*, **39**, (1961), 1253–1265.

2.* R.F.W. Bader, ''Vibrationally Induced Perturbations in Molecular Electron Distributions'', *Canadian Journal of Chemistry*, **40**, (1962), 1164–1175.

3. R.F.W. Bader and G.A. Jones, ''The Electron Density Distributions in Hydride Molecules, I, The Water Molecule'', *Canadian Journal of Chemistry*, **41**, (1963), 586–606.

4. R.F.W. Bader and G.A. Jones, ''The Electron Density Distribution in Hydride Molecules, II, The Ammonia Molecule'', *Journal of Chemical Physics*, **38**, (1963), 2791–2802.

5. R.F.W. Bader and G.A. Jones, ''The Electron Density Distributions in Hydride Molecules, III, The Hydrogen Fluoride Molecule'', *Canadian Journal of Chemistry*, **41**, (1963), 2251–2264.

6. R.F.W. Bader, ''Binding Regions in Polyatomic Molecules and Electron Density Distributions'', *Journal of the American Chemical Society*, **86**, (1964), 5070–5075.

7. R.F.W. Bader, W.H. Henneker and P.E. Cade, ''Molecular Charge Distributions and Chemical Binding'', *Journal of Chemical Physics*, **46**, (1967), 3341–3363.

8. R.F.W. Bader, I. Keaveny and P.E. Cade, ''Molecular Charge Distributions and Chemical Binding II. First-Row Diatomic Hydrides'', *Journal of Chemical Physics*, **47**, (1967), 3381–3402.

9. R.F.W. Bader and A.K. Chandra, ''A View of Bond Formation in Terms of Molecular Charge Distributions'', *Canadian Journal of Chemistry*, **46**, (1968), 953–966.

10. R.F.W. Bader and A.D. Bandrauk, ''Molecular Charge Distributions and Chemical Binding III. The Isoelectronic Series N_2, CO, BF and C_2, BeO, LiF'', *Journal of Chemical Physics*, **49**, (1968), 1653–1665.

11. R.F.W. Bader and A.D. Bandrauk, ''Relaxation of the Molecular Charge Distribution and the Vibrational Force Constant'', *Journal of Chemical Physics*, **49**, (1968), 1666–1675.

12. R.F.W. Bader and H.J.T. Preston, ''The Kinetic Energy of Molecular Charge Distributions and Molecular Stability'', *International Journal of Quantum Chemistry*, **3**, (1969), 327–347.

13. R.F.W. Bader, I. Keaveny and G. Runtz, ''Polarizations of Atomic and Molecular Charge Distributions'', *Canadian Journal of Chemistry*, **47**, (1969), 2308–2311.

14. R.F.W. Bader, P.E. Cade, W.H. Henneker and I. Keaveny, ''Molecular Charge Distributions and Chemical Binding IV. The Second-Row Diatomic Hydrides, AH'', *Journal of Chemical Physics*, **50**, (1969), 5313–5333.

15. R.F.W. Bader and J.L. Ginsburg, ''Relaxations of Molecular Charge Distributions and the Vibrational Force Constants in Diatomic Hydrides'', *Canadian Journal of Chemistry*, **47**, (1969), 3061–3074.

*This paper presents an early formulation of the symmetry rules predicting the outcome of unimolecular and bimolecular reactions. Kenichi Fukui describes this paper as ''the important theory of Bader'' in his 1981 Nobel Lecture in the paragraph he devotes to ''names which are worthy of special mention''.

or an extended system (Eq. (33) of Chapter 1). This result, termed the "atomic virial theorem", in its turn led to the definition of the energy of an atom within a larger system, for example a molecule or a crystal, i.e. an "atomic energy". The energy of an atom in a molecule, a very desirable quantity, remained totally elusive until the discovery of the atomic virial theorem, because this energy must, for example, include contributions from the nuclear–nuclear repulsion energy, contributions which are not trivial to partition on an atom-by-atom basis (Chapters 1 and 3).

Bader's early studies of molecular electron density distributions (Table 1) coincided with the ground-breaking formulation of modern density functional theory (DFT) [16] in 1964 and 1965 by Walter Kohn and his co-workers [17, 18]. Contemporary DFT functionals (for example those developed by Axel Becke [19–22], and Lee et al. [23]) are capable of achieving chemical accuracy and of producing electron-density maps of unprecedented quality, and rapidly.

The advent of DFT, the spectacular increase in the power of computers, and algorithmic advances all led to an explosive growth in the number of studies applying the quantum theory of atoms in molecules to a very wide range of problems (as will be seen in this book) from solid-state physics; to the science of materials; to surface science; to X-ray analyses; to organic, physical–organic, organometallic, and inorganic chemistry; and to biochemistry and drug design. Accurate calculated (and experimental) electron-density maps of larger and larger systems are now routinely computed and analyzed using the QTAIM.

The theory has also benefited significantly from parallel advances in accurate X-ray crystallography. The development of multipolar refinement techniques, pioneered by Hansen and Coppens [24–26], coupled with low-temperature data collection and ever-more sensitive CCD detectors, has enabled crystallographers, for the first time, to obtain high-resolution experimental electron-density maps of quality sufficient to capture the fine details of the electron density in the bonding regions between atoms. Nowadays, crystallographers rely routinely on QTAIM to decode the wealth of chemical information contained in accurate experimental electron-density maps, bringing crystallography and chemical theory closer than ever before (see, for example, Refs [25–28] and the literature cited therein).

Bader's landmark book [2], which includes (but is more than) an authoritative review of the theory up to 1990, sets forth the development and principles of this theory and explains how the atoms of experiment arise naturally from the laws of quantum mechanics. Since 1990 the field of QTAIM has grown dramatically both conceptually and in terms of the volume of publications and citations, a growth that has been reflected in several reviews (see, for example, Refs [3–6, 25–30]). In 1996, a special issue of the *Canadian Journal of Chemistry* was dedicated to Richard Bader on the occasion of his 65th birthday [31]. The objective of this book is to cover the developments in this field since the publication of Bader's book.

QTAIM is rigorous, beautiful, and powerful. It provides a unifying thread of physical insight in chemistry, which explains its popularity. The breadth of QTAIM and its applications renders a comprehensive treatment of all its ramifi-

cations impossible in a book of this size. We have therefore *sampled* research in QTAIM by extending invitations to a necessarily incomplete group of world-leading researchers to review their respective contributions to the field. This has resulted in a volume written by fifty authors representing thirteen countries in five different continents (a list of contributors is given below). Despite this impressive list of contributors, we could not possibly have invited all the leaders of the field, unavoidable omissions for which we do apologize. These omissions, however, do not diminish the value of the phenomenal cross-section and depth of current fundamental and applied research in QTAIM that has been captured in this book.

As the editors of this book it is with considerable humility that we start with our own introductory chapter. The *only* reason for this choice is to facilitate the reading of the remainder of the book by introducing the basic concepts and terminology. The order of the other parts and chapters is purely and exclusively based on what we think is their logical order. All chapters, including our own, have been carefully refereed by at least three independent reviewers and were all revised and corrected before final acceptance.

The book is divided into five parts. The introductory chapter is followed by Part II which concentrates on the fundamental advances in the theory itself. Part II reviews the rapid development in the applications of the QTAIM to periodic systems (solid state and surfaces). Part III focuses on developments resulting from the synergy between experimental highly accurate X-ray crystallography and the QTAIM, with particular emphasis on the electron density of large biological molecules. Part VI deals with the wide diversity of applications of the QTAIM in organic, physical organic, and organometallic chemistry, and reviews the characterization of conventional and non-conventional chemical bonding. Part V reports on important developments in the use of QTAIM in the modeling of biological molecules and drug design.

We would like to thank each one of the authors individually for his or her invaluable contribution to this volume, and we thank Professor Axel D. Becke for writing the Foreword. We are much indebted to our publisher, Wiley–VCH, and its staff for their continual support, professionalism, and invaluable help, with special thanks to Nele Denzau, Dr. Tim Kersebohm, Dr. Romy Kirsten, Claudia Nussbeck, Dr. Martin Ottmar, Dr. Gudrun Walter, and Dr. Waltraud Wüst. We are particularly grateful for the care, rigor, and effort of our peer-reviewers (many of whom are also among the authors of chapters in this book): Professor Richard F.W. Bader, Dr. Miguel Blanco, Professor Curt M. Breneman, Dr. Clémence Corminboeuf, Professor Katherine V. Darvesh, Dr. Jason R. Dwyer, Dr. Carlo Gatti, Professor Kathleen M. Gough, Professor Sławomir J. Grabowski, Professor George L. Heard, Professor Jesús Hernández-Trujillo, Dr. Sian T. Howard, Professor Claude Lecomte, Professor Victor Luaña, Professor Peter Luger, Dr. Piero Macchi, Professor Preston J. MacDougall, Professor Louis J. Massa, Professor Angel Martín Pendás, Dr. James A. Platts, Dr. Paul L.A. Popelier, Dr. Kathy N. Robertson, Professor Bernard Silvi, Professor Vladimir G. Tsirelson, and Dr. Elizabeth A. Zhurova – we thank them all.

It is with delight that we dedicate this work to Professor Richard F.W. Bader on the occasion of his 75th birthday.

Halifax, October 2006 *Chérif F. Matta and Russell J. Boyd*

References

1 R.F.W. Bader, G.A. Jones; The electron density distribution in hydride molecules. The ammonia molecule. *J. Chem. Phys.* **1963**, *38*, 2791–2802.

2 R.F.W. Bader; *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.

3 R.F.W. Bader; The quantum mechanical basis of conceptual chemistry. *Monatsh. Chem.* **2005**, *136*, 819–854.

4 R.F.W. Bader; 1997 Polanyi Award Lecture: Why are there atoms in chemistry? *Can. J. Chem.* **1998**, *76*, 973–988.

5 R.F.W. Bader, in: P.v.-R. Schleyer (Ed.); *Encyclopedia of Computational Chemistry*; John Wiley and Sons: Chichester, UK, 1998, pp 64–86.

6 P.L.A. Popelier; *Atoms in Molecules: An Introduction*; Prentice Hall: London, 2000.

7 R.F.W. Bader, P.L.A. Popelier, T.A. Keith; Theoretical definition of a functional group and the molecular orbital paradigm. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 620–631.

8 C.F. Matta, R.F.W. Bader; An experimentalist's reply to "What is an atom in a molecule?". *J. Phys. Chem. A* **2006**, *110*, 6365–6371.

9 R.F.W. Bader; Principle of stationary action and the definition of a proper open system. *Phys. Rev. B* **1994**, *49*, 13348–13356.

10 R.F.W. Bader; Quantum topology of molecular charge distributions. III. The mechanics of an atom in a molecule. *J. Chem. Phys.* **1980**, *73*, 2871–2883.

11 P.F. Zou, R.F.W. Bader; Variational principle and path integrals for atoms in molecules. *Int. J. Quantum Chem.* **1992**, *43*, 677–699.

12 R.F.W. Bader, P.M. Beddall; Virial field relationship for molecular charge distributions and the spatial partitioning of molecular properties. *J. Chem. Phys.* **1972**, *56*, 3320–3328.

13 R.F.W. Bader, P.M. Beddall, J. Jr. Peslak; Theoretical development of a virial relationship for spatially defined fragments of molecular systems. *J. Chem. Phys.* **1973**, *58*, 557–566.

14 S. Srebrenik, R.F.W. Bader; Towards the development of the quantum mechanics of a subspace. *J. Chem. Phys.* **1975**, *63*, 3945–3961.

15 S. Srebrenik, R.F.W. Bader, T.T. Nguyen-Dang; Subspace quantum mechanics and the variational principle. *J. Chem. Phys.* **1978**, *68*, 3667–3679.

16 R.G. Parr, W. Yang; *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989.

17 P. Hohenberg, W. Kohn; Inhomogeneous electron gas. *Phys. Rev. B* **1964**, *136*, 864–871.

18 W. Kohn, L.J. Sham; Self consistent equations including exchange and correlation effects. *Phys. Rev. A* **1965**, *140 (4A)*, 1133–1138.

19 A. Becke; Density-functional thermochemistry.1. The effect of the exchange-only gradient correction. *J. Chem. Phys.* **1992**, *96*, 2155–2160.

20 A. Becke; Density-functional thermochemistry.2. The effect of the Perdew-Wang generalized-gradient correlation correction. *J. Chem. Phys.* **1992**, *97*, 9173–9177.

21 A. Becke; A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

22 A. Becke; Density-functional thermochemistry.3. The role of exact

exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

23 C. Lee, W. Yang, R. Parr; Development of the Colle–Salvetti correlation-energy formula into a functional of the electron-density. *Phys. Rev. B* **1988**, *37*, 785–789.

24 N.K. Hansen, P. Coppens; Testing aspherical atom refinement on small molecules data sets. *Acta Cryst.* **1978**, *A34*, 909–921.

25 P. Coppens; *X-ray Charge Densities and Chemical Bonding*; Oxford University Press, Inc.: New York, 1997.

26 T.S. Koritsanszky, P. Coppens; Chemical applications of X-ray charge-density analysis. *Chem. Rev.* **2001**, *101*, 1583–1628.

27 C. Gatti; Chemical bonding in crystals: New directions. *Z. Kristallogr.* **2005**, *220*, 399–457.

28 P. Macchi, A. Angelo Sironi; Chemical bonding in transition metal carbonyl clusters: Complementary analysis of theoretical and experimental electron densities. *Coord. Chem. Rev.* **2003**, *238–239*, 383–412.

29 R.F.W. Bader; Can there be more than a single definition of an atom in a molecule? *Can. J. Chem.* **1999**, *77*, 86–93.

30 R.J. Gillespie, P.L.A. Popelier; *Molecular Geometry and Chemical Bonding: From Lewis to Electron Densities*; Oxford University Press: New York, 2001.

31 R.J. Boyd, R.A. McLelland, N.H. Werstiuk (Guest Editors); *Can. J. Chem. (Special issue dedicated to Professor Richard F.W. Bader)* **1996**, *74*(6).

List of Abbreviations Appearing in this Volume

This list is compiled from the abbreviations appearing in this book, and is not, therefore, a comprehensive compilation of abbreviations in quantum chemistry. This list of abbreviations is, furthermore, naturally slanted more toward QTAIM by the inclusion of the abbreviations appearing in this work, for the convenience of the reader.

(Excellent resources for abbreviations in quantum chemistry include: [1] R.D. Brown, J.E. Bows, R. Hilderbrandt, K. Lim, I.M. Mills, E. Nikitin, M.H. Palmer; Acronyms used in Theoretical Chemistry, *Pure & Appl. Chem.*, **68**, 387–456, (1996); and [2] Young, D.; Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems, Wiley–Interscience, New York (2001), pp. 360–370.)

A	Adenine
AD	Atomic dipole
ADMET	Absorption, distribution, metabolism, excretion, and toxicity
AIM	Atoms in molecules (or sometimes it implies an atom in a molecule)
aiPI	ab initio perturbed ion (method)
ANO	Atomic natural orbital(s)
AO	Atomic orbital
ASE	Aromatic stabilization energy
au or a.u.	Atomic unit
aug-cc-pVTZ	Dunning's valence triple zeta correlation consistent basis set augmented with diffuse functions
B3LYP	Becke parameter 3, Lee, Yang, and Parr (a gradient-corrected DFT functional)
B96	Becke 1996 gradient-corrected functional
BCP or bcp	Bond critical point
BNP	Bare nuclear potential
BO	Born-Oppenheimer
BO	Bond order
BP	Bond path
BPL	Bond path-length

BSSE	Basis set superposition error
C	Cytosine
CAHB	Charge-assisted hydrogen bond(ing)
CAHB(−)	Negative-charge-assisted hydrogen bond(ing)
CAHB(+)	Positive-charge-assisted hydrogen bond(ing)
CASSCF	Complete active space self-consistent field (theory/method)
cc	Complex conjugate
CC	Coupled-cluster (correlated electronic structure method)
CC	Charge concentration
CCD	Charge-coupled device (x-ray detector)
CCP or ccp	Cage critical point
CCSD	Coupled-cluster method with single and double excitations
CCSD(T)	Coupled-cluster method with single and double excitations with singles/triples coupling term
CD	Degree of covalence
CD(s)	Charge depletion(s)
CG	Crystal geometry (as opposed to gas-phase geometry)
CI	Configuration interaction
CISD	Configuration interaction with single and double excitations
CKRK	Cash–Karp–Runge–Kutta (method)
CORR	Correlation energy
Cp	Cyclopentadienyl ring
CP or cp	Critical point
CPHF	Coupled-perturbed Hartree–Fock (method)
CT	Charge transfer
CUS	Coordinatively unsaturated site(s)
c.o.m.	Center of mass
DCBS	Dimer-centered basis set
DCD	Dewas-Chatt-Duncanson (donor-acceptor complexes)
DFT	Density functional theory
DHB	Dihydrogen bond
DI	Delocalization index
DMACB	3,4-bis(Dimethylamino)-3-cyclobutene-1,2-dione
DMSDA	Difference of the mean-square displacement amplitude
DNA	Deoxyribonucleic acid
DOS	Density of (electronic) states
ECHB	Electrostatic–covalent hydrogen bond (model)
ECP	Effective core potential
ED	Electron density
EDD	Electron density distribution(s)
EF	Eigenvalue following (method)
ELF	Electron localization function (Becke and Edgecombe)
EP	Electrostatic potential
ESP	Electrostatic potential
EXAFS	Extended X-ray absorption fine structure

FCC	Face-centered cubic (lattice)
FLU	Aromatic fluctuation index (of Solà et al.)
fpLAPW	Full potential linearized augmented plane wave (formalism)
FT	Fourier transform(ation)
G	Guanine
GGA	Generalized gradient approximation
GVB	Generalized valence bond (method)
hAR	Human aldose reductase
HB	Hydrogen bond(ing)
HDS	Hydrodesulfurization
HF	Hartree–Fock (not to be confused with hydrogen fluoride)
HK	Hohenberg and Kohn
HOMA	Harmonic oscillator model of aromaticity (index)
HOMO	Highest occupied molecular orbital
HPLC	High-pressure (or high-performance) liquid chromatography
i-VSCC	Inner-valence shell charge concentration
IAM	Independent atom model
IAS	Inter-atomic surface
IGAIM	Individual gauges for atoms in molecules
IHB	Isolated hydrogen bond
IRC	Intrinsic reaction coordinate(s)
IUCr	International Union of Crystallography
KE	Kinetic energy
KS	Kohn–Sham
LCAO	Linear combination of atomic orbitals
LDA	Local density approximation
LEED	Low-energy electron diffraction
LFT	Ligand field theory
LI	Localization index
LOCC	Ligand-opposed charge concentration
LS	Local source
LUMO	Lowest unoccupied molecular orbital
L.H.S.	Left-hand side (of an equation)
MD	Molecular dynamics
MEM	Maximum entropy method
MO	Molecular orbital
MP*n*	*n*th order Møller–Plesset perturbation theory
MQSM	Molecular quantum similarity measures
MRCI	Multireference configuration interaction
***m*RNA**	Messenger-RNA
***n*-MR**	*n*-Membered ring, e.g. five-membered ring
NA	Nuclear attractor
NADP+	Nicotinamide adenine dinucleotide phosphate
NBCC	Nonbonded charge concentration
NBM	Nonbonded maximum (or maxima)

NBO	Natural bond orbital(s) (method)
NCP or ncp	Nuclear critical point
NICS	(Schleyer's) nucleus-independent chemical shift
NICS(0)	(Schleyer's) nucleus-independent chemical shift in the plane of an aromatic ring at its geometrical center
NICS(*n*)	(Schleyer's) nucleus-independent chemical shift *n* Å above the plane of an aromatic ring at its geometrical center
NMR	Nuclear magnetic resonance (spectroscopy)
NNA	Non-nuclear attractor
NNA	Nearest-neighbor approximation
NNM	Non-nuclear maximum (or maxima)
NPA	Natural population analysis
NR	Newton–Raphson (method/technique)
ODM	One-electron density matrix
OG	(Gas-phase) optimized geometry (as opposed to crystal geometry)
PAHB	Polarization-assisted hydrogen bond(ing)
PAH(s)	Polycyclic aromatic hydrocarbon(s)
PASA	Promolecular atom shell approximation
PBH(s)	Polybenzenoid hydrocarbon(s)
PDI	*para*-Delocalization index (of Solà et al.)
PE	Potential energy
PES	Potential energy surface(s)
PEST	Property-encoded surface translator
PH	Poincaré-Hopf relationship
ppm	Part per million
Pyr	Pyridine
QCISD	Quadratic configuration interaction with single and double excitations
QCISD(T)	Quadratic configuration interaction with single and double excitations (with correction for triple excitations)
QCT	Quantum chemical topology
QSAR	Quantitative structure–activity relationship(s)
QSPR	Quantitative structure–property relationship(s)
QTAIM	Quantum theory of atoms in molecules
QTAMC	Quantum theory of atoms in molecules and crystals
RAHB	Resonance-assisted hydrogen bond(ing)
RCP or rcp	Ring critical point
RECON	(Rapid TAE) reconstruction (of molecular charge densities)
RHF	Restricted Hartree–Fock (method)
RNA	Ribonucleic acid
RSS	Root summed squares
R.H.S.	Right-hand side (of an equation)
SCF	Self-consistent field
SCVS	Self-consistent virial scaling
SES	Sodium electrosodalite

SF	Source function
SI	Surface integral
SIEP	Surface integral of the electrostatic potential
SIG	Surface integral of the G kinetic energy density
SIK	Surface integral of the K kinetic energy density
SSHB	Short strong hydrogen bond(ing)
std	Standard deviation
STM	Scanning–tunneling microscope (or microscopy)
T	Thymine
TAE	(Breneman's) transferable atom equivalent
TMS	Transition metal sulfide(s)
TMS	Tetramethylsilane
U	Uracil
UHF	Unrestricted Hartree–Fock (method)
VHTS	Virtual high-throughput screening
VSCC	Valence shell charge concentration
VSCD	Valence shell charge depletion
VSEPR	Valence-shell electron-pair repulsion model of molecular geometry
XRD	X-Ray diffraction
ZFS	Zero-flux surface

List of Contributors

Dr. Ibon Alkorta
Instituto de Química Médica
(CSIC)
Juan de la Cierva 3
28006 Madrid
Spain
ibon@iqm.csic.es

Dr. Yosslen Aray
Centro de Quimica, IVIC
Apartado 21827
Caracas 1020 A
Venezuela
yaray@mail.ivic.ve

Prof. Richard F.W. Bader
Department of Chemistry
McMaster University
1280 Main Street West
Hamilton, ON, L4S 4M1
Canada
bader@mcmaster.ca

Prof. Axel D. Becke
Department of Chemistry
Dalhousie University
Halifax, NS, B3H 4J3
Canada
axel.becke@dal.ca

Dr. Miguel A. Blanco
Departamento de Química Física
y Analítica
Universidad de Oviedo
33006 Oviedo
Spain
miguel@carbono.quimica.uniovi.es

Prof. Russell J. Boyd
Department of Chemistry
Dalhousie University
Halifax, NS, B3H 4J3
Canada
russell.boyd@dal.ca

Prof. Curt M. Breneman
Department of Chemistry and
Chemical Biology
Rensselaer Polytechnic Institute
110 8th St
Troy, NY 12180-3590
USA
brenec@rpi.edu

Prof. Fernando Cortés-Guzmán
Departamento de Química Orgánica
Facultad de Química
Universidad Nacional Autónoma de
México
Circuito Escolar, Ciudad Universitaria
Coyoacán, México D.F. 04510
fercor@servidor.unam.mx

Dr. Aurora Costales
Departamento de Química Física
y Analítica
Universidad de Oviedo
33006 Oviedo
Spain
yoyi@carbono.quimica.uniovi.es

Prof. Gabriel Cuevas
Departamento de Fisicoquímica
Instituto de Química
Universidad Nacional Autónoma
de México
Circuito Escolar, Ciudad
Universitaria
Coyoacán, México D.F. 04510
México
gecgb@servidor.unam

Dr. Richard Dawes
Department of Chemistry
University of Missouri-Columbia
Columbia, MO 65211
USA
dawesr@missouri.edu

Dr. Birger Dittrich
School of Biomedical & Chemical
Sciences
The University of Western
Australia
35 Stirling Highway
Crawley WA 6009
Australia
birger@chemie.fu-berlin.de
birger@cyllene.uwa.edu.au

Dr. Jason R. Dwyer
Max-Born-Institut für
Nichtlineare Optik und
Kurzzeitspektroskopie
Max-Born-Strasse 2A
12489 Berlin
Germany
dwyer@mbi-berlin.de

Prof. José Elguero
Instituto de Química Médica (CSIC)
Juan de la Cierva 3
28006 Madrid
Spain
iqmbe17@iqm.csic.es

Prof. Enrique Espinosa
Laboratoire de Cristallographie et
Modélisation des Matériaux Minéraux
et Biologiques
Faculté des Sciences et Techniques
Université Henri Poincaré
Nancy 1
BP 239, 54506 Vandoeuvre-lés-
Nancy, Cedex
France
Enrique.Espinosa@lcm3b.uhp-nancy.fr

Dr. Carlo Gatti
CNR-ISTM Istituto di Scienze e
Tecnologie Molecolari
Via Golgi 19
20133 Milano
Italy
c.gatti@istm.cnr.it

Prof. Ronald J. Gillespie
Department of Chemistry
McMaster University
1280 Main Street West
Hamilton, ON, L4S 4M1
Canada
ronald.gillespie@sympatico.ca
gillespi@mcmaster.ca

Prof. Kathleen M. Gough
Department of Chemistry
University of Manitoba
Winnipeg, MB, R3T 2N2
Canada
kmgough@cc.umanitoba.ca

Prof. Sławomir J. Grabowski
Department of Physics and Chemistry
University of Łódź
ul.Pomorska 149/153
90-236 Łódź
Poland
slagra@ccmsi.us

Dr. Benoît Guillot
Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques
LCM3B (UMR UHP-CNRS 7036)
Faculté des Sciences et Techniques
Université Henri Poincaré
BP 239, Boulevard des Aiguillettes
54506 Vandoeuvre-lès-Nancy CEDEX
France
benoit.guillot@lcm3b.uhp-nancy.fr

Dr. Christopher E. Henze
MS T27-A
NASA Advanced Supercomputing Division
NASA Ames Research Center
Moffett Field, CA 94035
USA
chenze@nas.nasa.gov

Prof. Jesús Hernández-Trujillo
Departamento de Física y Química Teórica
Facultad de Química
Universidad Nacional Autónoma de México
Circuito Escolar, Ciudad Universitaria
Coyoacán, México D.F.
México 04510
jesus.hernandez@correo.unam.mx

Dr. Christian Jelsch
Laboratoire de Cristallographie et Modelisation des Matériaux Minéraux et Biologiques
Univérsité Henri Poincaré
LCM3B (UMR UHP-CNRS 7036)
Faculté des Sciences et Techniques
Université Henri Poincaré
BP 239, Boulevard des Aiguillettes
54506 Vandoeuvre-lès-Nancy CEDEX
France
christian.jelsch@lcm3b.uhp-nancy.fr

Dr. Todd A. Keith
Department of Chemistry
University of Missouri – Kansas City
5100 Rockhill Road
Kansas City, MO 64110
USA
keithta@umkc.edu

Dr. Angélique Lagoutte
Laboratoire de Cristallographie et Modelisation des Matériaux Minéraux et Biologiques
Univérsité Henri Poincaré
LCM3B (UMR UHP-CNRS 7036)
Faculté des Sciences et Techniques
Université Henri Poincaré
BP 239, Boulevard des Aiguillettes
54506 Vandoeuvre-lès-Nancy CEDEX
France
angelique.lagoutte@lcm3b.uhp-nancy.fr

Prof. Claude Lecomte
Laboratoire de Cristallographie et Modelisation des Matériaux Minéraux et Biologiques
Univérsité Henri Poincaré
LCM3B (UMR UHP-CNRS 7036)
Faculté des Sciences et Techniques
Université Henri Poincaré
BP 239, Boulevard des Aiguillettes
54506 Vandoeuvre-lès-Nancy CEDEX
France
claude.lecomte@lcm3b.uhp-nancy.fr

Prof. Laurence Leherte
Laboratoire de Physico-Clinic Informatique (PCI)
Facultés Universitaires Notre-Dame de la Paix
Rue de Bruxelles 61
5000 Namur
Belgium
laurence.leherte@fundp.ac.be

Prof. Víctor Luaña
Departamento de Química Física y Analítica
Universidad de Oviedo
33006 Oviedo
Spain
victor@carbono.quimica.uniovi.es

Prof. Peter Luger
Institute for Chemistry/ Crystallography
Free University of Berlin
Takustr. 6
14195 Berlin
Germany
luger@chemie.fu-berlin.de

Dr. Piero Macchi
Dipartimento di Chimica Strutturale e Stereochimica Inorganica
Università di Milano
Via Venezian 21
20133 Milano
Italy

and

CNR – Istituto di Scienze e Tecnologie Molecolari
Via Golgi 19
Milano
Italy
piero.macchi@istm.cnr.it

Prof. Preston J. MacDougall
Department of Chemistry
Middle Tennessee State University
1301 E. Main St.
Murfreesboro, TN 37132
USA
pmacdougall@mtsu.edu
preston.macdougall@gmail.com

Dr. Ignasi Mata
Institut de Ciència de Materials de Barcelona (CSIC)
Campus UAB
08193 Bellaterra
Spain
imata@icmab.es

Dr. Eduard Matito
Institut de Química Computacional
Departament de Química
Universitat de Girona
Campus de Montilivi
17071 Girona
Catalonia
Spain
eduard@iqc.udg.es

Prof. Chérif F. Matta
Department of Chemistry & Physics
Mount Saint Vincent University
Halifax, NS, B3M 2J6
Canada

and

Department of Chemistry
Dalhousie University
Halifax, NS, B3H 4J3
Canada
cherif.matta@msvu.ca
quantumjazz@hotmail.com

Prof. Elies Molins
Institut de Ciència de Materials
de Barcelona (CSIC)
Campus UAB
08193 Bellaterra
Spain
elies.molins@icmab.es

Dr. Paula Mori-Sánchez
Departamento de Química Física
y Analítica
Universidad de Oviedo
33006 Oviedo
Spain
paula@carbono.quimica.uniovi.es

Prof. Angel Martín Pendás
Departamento de Química Física
y Analítica
Universidad de Oviedo
33006 Oviedo
Spain
angel@carbono.quimica.uniovi.es

Dr. Virginie Pichon-Pesme
Laboratoire de Cristallographie et
Modélisation des Matériaux
Minéraux et Biologiques
LCM3B (UMR UHP-CNRS 7036)
Faculté des Sciences et
Techniques
Université Henri Poincaré
BP 239, Boulevard des
Aiguillettes
54506 Vandoeuvre-lès-Nancy
CEDEX
France
virginie.pichon@lcm3b.uhp-
nancy.fr

Dr. Jordi Poater
Afdeling Theoretische Chemie
Scheikundig Laboratorium der Vrije
Universiteit
De Boelelaan 1083
NL-1081 HV Amsterdam
The Netherlands
J.Poater@few.vu.nl

Dr. Paul L. A. Popelier
School of Chemistry
University of Manchester
Faraday Building, North Campus
Manchester M60 1QD
UK
paul.popelier@manchester.ac.uk

Dr. Michel Rafat
School of Chemistry
University of Manchester
Faraday Building, North Campus
Manchester M60 1QD
UK
m-rafat@inbox.com

Dr. Jesus Rodríguez
Centro de Quimica, IVIC
Apartado 21827
Caracas 1020 A
Venezuela
jrodrigu@ivic.ve

Prof. Bernard Silvi
Laboratoire de Chimie Théorique
UMR (CNRS) 7616
Université Pierre et Marie Curie
3, Rue Galilée
94200 Ivry sur Seine
France
silvi@lct.jussieu.fr

Prof. Angelo Sironi
Dipartimento di Chimica
Strutturale e Stereochimica
Inorganica
Università di Milano
Via Venezian 21
20133 Milano
Italy

and

CNR – Istituto di Scienze e
Tecnologie Molecolari
Via Golgi 19
Milano
Italy
angelo.sironi@istm.cnr.it

Prof. Miquel Solà
Institut de Química
Computacional
Departament de Química
Universitat de Girona
Campus de Montilivi
17071 Girona
Catalonia
Spain
miquel.sola@udg.es

Dr. Nagamani Sukumar
Department of Chemistry and
Biological Chemistry
Center for Biotechnology and
Interdisciplinary Studies
Rensselaer Polytechnic Institute
110 8th St.
Troy, NY 12180-3590
USA
nagams@rpi.edu

Prof. Vladimir G. Tsirelson
Quantum Chemistry Department
Mendeleev University of Chemical
Technology
Miusskaya Sq. 9
Moscow 125047
Russia
tsirel@muctr.edu.ru
vtsirelson@yandex.ru

Dr. David Vega
Departamento de Química
Facultad Experimental de Ciencias y
Tecnología (FACYT)
Universidad de Carabobo
Valencia, Edo. Carabobo
Venezuela
dvega@uc.edu.ve

Prof. Daniel P. Vercauteren
Laboratoire de Physico-Clinic
Informatique (PCI)
Facultés Universitaires Notre-Dame de
la Paix (FUNDP)
Rue de Bruxelles 61
5000 Namur
Belgium
daniel.vercauteren@fundp.ac.be

Dr. Tammy L. Welshman
Department of Chemistry
University of Manitoba
Winnipeg, MB, R3T 2N2
Canada
welshman@cc.umanitoba.ca

1
An Introduction to the Quantum Theory of Atoms in Molecules

Chérif F. Matta and Russell J. Boyd

1.1
Introduction

The observation that some properties attributed to atoms and functional groups are transferable from one molecule to another has played a key role in the development of chemistry. This observation provides a basis for group additivity schemes and is exemplified by the constancy of group contributions to thermodynamic and spectroscopic properties. But what is the electronic basis of this empirical transferability? The quantum theory of atoms in molecules (QTAIM) [1], developed by Professor Richard F. W. Bader and his coworkers, relies on quantum observables such as the electron density $\rho(\mathbf{r})$ and energy densities to answer such a question. Other important (related) questions addressed by QTAIM include:

- What is an atom in a molecule or a crystal?
- How can an atom or a group of atoms be transferable sometimes in *very* different external potentials?
- Can one define bonding in molecules unambiguously especially in borderline cases?

This chapter contains a summary of some of the main concepts of QTAIM. A more comprehensive and mathematically elegant treatment can be found in Bader's book [1].

(Often in this chapter, the word ''molecule'' includes extended systems such as polymers, weakly bonded molecular complexes, and molecular and ionic crystals, in addition to its more traditional meaning of a single, finite, isolated chemically bonded group of atoms. It will be clear from the context when this term is used in its traditional or in its larger sense.)

1.2
The Topology of the Electron Density

The topology of the electron density is dominated by the attractive forces of the nuclei imparting it with its principal topological feature – a substantial local max-

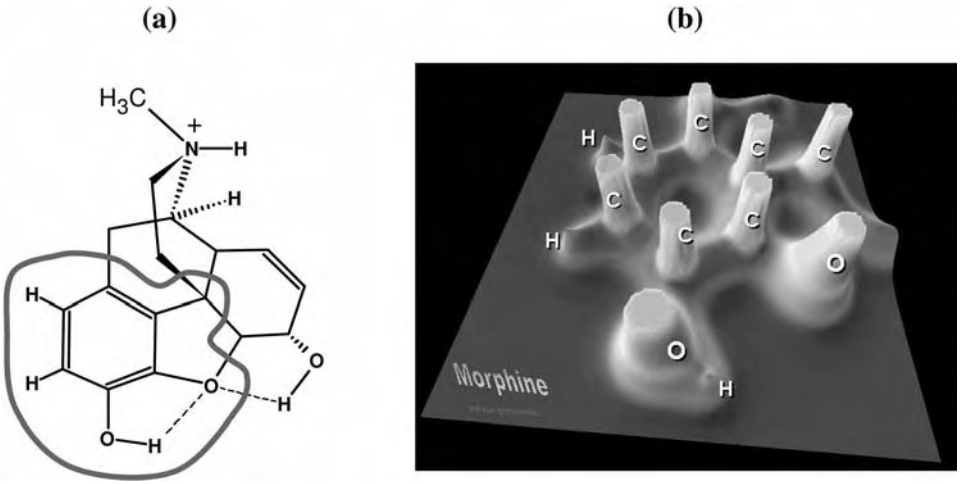


Fig. 1.1 (a) The molecular structure of the morphine molecule with an indication of the region shown in the relief map in (b). (b) A relief map representation of the electron density in the plane of the aromatic ring showing marked maxima at the positions of the carbon and oxygen nuclei (truncated at $\rho(\mathbf{r}) = 1.0$ au) and much smaller peaks at the position of the hydrogen nuclei.

imum at the position of each nucleus. A consequence of the dominance of nuclear maxima in the electron density distribution is the association of an atom with a region of space the boundaries of which are determined by the balance in the forces the neighboring nuclei exert on the electrons. Figure 1.1b is a relief map of the electron density of the phenolic region of the morphine molecule, in the plane of the aromatic ring, showing the maxima at the C, O, and H nuclei.

A "critical point" (CP) in the electron density is a point in space at which the first derivatives of the density vanish, i.e.:

$$\nabla\rho = \mathbf{i}\frac{d\rho}{dx} + \mathbf{j}\frac{d\rho}{dy} + \mathbf{k}\frac{d\rho}{dz} \rightarrow \begin{cases} = \vec{\mathbf{0}} & \text{(At critical points and at } \infty\text{)} \\ \text{Generally} \neq \vec{\mathbf{0}} & \text{(At all other points)} \end{cases} \quad (1)$$

where the zero vector signifies that each individual derivative in the gradient operator, ∇, is zero and not just their sum. The gradient of a scalar function such as $\rho(\mathbf{r})$ (Eq. 1) at a point in space is a vector pointing in the direction in which $\rho(\mathbf{r})$ undergoes the greatest rate of increase and having a magnitude equal to the rate of increase in that direction. The maximum at the position of a nucleus constitutes one type of CP, namely, a nuclear critical point (NCP). (The neglect of the finite size of atomic nuclei in quantum chemical calculations, an exceptionally good approximation, results in cusps in the potential and in the electron density $\rho(\mathrm{r})$ at the position of the nuclei. Because of this cusp, the derivatives of the electron

density at the position of a nucleus are not defined and so, in a formal mathematical sense, this position is not a true critical point. The nuclear maxima behave *topologically* as critical points, however.)

One can discriminate between a local minimum, a local maximum, or a saddle point by considering the second derivatives, the elements of the tensor $\nabla\nabla\rho$. There are nine second derivatives of $\rho(\mathbf{r})$ that can be arranged in the so-called "Hessian matrix", which when evaluated at a CP located at $\mathbf{r}_c$ is written:

$$\mathbf{A}(\mathbf{r}_c) = \begin{pmatrix} \frac{\partial^2\rho}{\partial x^2} & \frac{\partial^2\rho}{\partial x\partial y} & \frac{\partial^2\rho}{\partial x\partial z} \\ \frac{\partial^2\rho}{\partial y\partial x} & \frac{\partial^2\rho}{\partial y^2} & \frac{\partial^2\rho}{\partial y\partial z} \\ \frac{\partial^2\rho}{\partial z\partial x} & \frac{\partial^2\rho}{\partial z\partial y} & \frac{\partial^2\rho}{\partial z^2} \end{pmatrix}_{\mathbf{r}=\mathbf{r}_c} . \tag{2}$$

The Hessian matrix can be diagonalized because it is real and symmetric. The diagonalization of $\mathbf{A}(\mathbf{r}_c)$ is equivalent to a rotation of the coordinate system $\mathbf{r}(x, y, z) \rightarrow \mathbf{r}(x', y', z')$ superimposing the new axes x', y', z' with the principal curvature axes of the critical point. The rotation of the coordinate system is accomplished via a unitary transformation, $\mathbf{r}' = \mathbf{r}\mathbf{U}$, where $\mathbf{U}$ is a unitary matrix constructed from a set of three eigenvalue equations $\mathbf{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$ $(i = 1, 2, 3)$ in which $\mathbf{u}_i$ is the ith column vector (eigenvector) in $\mathbf{U}$. A similarity transformation $\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{\Lambda}$ transforms the Hessian into its diagonal form, which is written explicitly as:

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{\partial^2\rho}{\partial x'^2} & 0 & 0 \\ 0 & \frac{\partial^2\rho}{\partial y'^2} & 0 \\ 0 & 0 & \frac{\partial^2\rho}{\partial z'^2} \end{pmatrix}_{\mathbf{r}'=\mathbf{r}_c} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \tag{3}$$

in which λ_1, λ_2, and λ_3 are the curvatures of the density with respect to the three principal axes x', y', z'.

An important property of the Hessian is that its trace is invariant to rotations of the coordinate system. The trace of the Hessian of the density is known as the Laplacian of the density $[\nabla^2\rho(\mathbf{r})]$ and, when $x = x'$, $y = y'$, and $z = z'$, is given by:

$$\nabla^2\rho(\mathbf{r}) = \nabla\cdot\nabla\rho(\mathbf{r}) = \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial x^2}}_{\lambda 1} + \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial y^2}}_{\lambda 2} + \underbrace{\frac{\partial^2\rho(\mathbf{r})}{\partial z^2}}_{\lambda 3} \tag{4}$$

where we have dropped the primes of the principal axes.

Critical points are classified according to their *rank* (ω) and *signature* (σ) and are symbolized by (ω, σ). The rank is the number of non-zero curvatures of ρ at the critical point. A critical point that has $\omega < 3$ is mathematically unstable and will vanish or bifurcate under small perturbations of the density caused by nuclear motion. The presence of such a CP (with a rank less than three) indicates a change in the topology of the density and, hence, a change in the molecular structure. For this reason, critical points with $\omega < 3$ are generally not found in equilibrium charge distributions and one nearly always finds $\omega = 3$. The signature is the algebraic sum of the signs of the curvatures, i.e. each of the three curvatures contributes ± 1 depending on whether it is a positive or negative curvature.

There are four types of stable critical points having three non-zero eigenvalues:

- $(3, -3)$ Three negative curvatures: ρ is a local maximum.
- $(3, -1)$ Two negative curvatures: ρ is a maximum in the plane defined by the corresponding eigenvectors but is a minimum along the third axis which is perpendicular to this plane.
- $(3, +1)$ Two positive curvatures: ρ is a minimum in the plane defined by the corresponding eigenvectors and a maximum along the third axis which is perpendicular to this plane.
- $(3, +3)$ Three curvatures are positive: ρ is a local minimum.

Each type of critical point described above is identified with an element of chemical structure: $(3, -3)$ *nuclear critical point* (NCP); $(3, -1)$ *bond critical point* (BCP); $(3, +1)$ *ring critical point* (RCP); and $(3, +3)$ *cage critical point* (CCP).

The number and type of critical points that can coexist in a molecule or crystal follow a strict topological relationship which states that:

$$n_{\mathrm{NCP}} - n_{\mathrm{BCP}} + n_{\mathrm{RCP}} - n_{\mathrm{CCP}} = \begin{cases} 1 & \text{(Isolated molecules)} \\ 0 & \text{(Infinite crystals)} \end{cases} \tag{5}$$

where n denotes the number of the subscripted type of CP. The first equality is known as the Poincaré–Hopf relationship (PH) [1] and applies for isolated finite systems such as a molecule, the second equality is known as the Morse equation and applies in cases of infinite periodic lattices [2]. The set $\{n_{\mathrm{NCP}}, n_{\mathrm{BCP}}, n_{\mathrm{RCP}}, n_{\mathrm{CCP}}\}$ for a given system is known as the "characteristic set".

Violation of Eq. (5) implies an inconsistent characteristic set, that a critical point has been missed, and that a further search for the missing critical point(s) is necessary. On the other hand, the satisfaction of this equation does not *prove* its completeness. For example, if we miss both a BCP and an RCP for a molecule, Eq. (5) becomes $n_{\mathrm{NCP}} - (n_{\mathrm{BCP}} - 1) + (n_{\mathrm{RCP}} - 1) - n_{\mathrm{CCP}} = 1$ which is clearly still valid [3]. The likelihood of missing both a BCP and a RCP is small, however, and, in practice, satisfaction of Eq. (5) is taken as a proof of the consistency *and* completeness of the characteristic set.

A ring critical point will always be found in the interior of a ring of chemically bonded atoms. When several rings are connected in a manner which encloses an

interstitial space, a cage critical point arises in the enclosed space. Figure 1.2 shows the molecular graph (the set of bond paths and critical points) of two molecules: (a) cubane, and (b) 4-methyl-1,12-difluoro[4]helicene. The bond path is a single line of maximum electron density linking the nuclei of two chemically-bonded atoms. (The bond path is discussed in more detail later in this chapter.) In cubane, the bond paths are arranged between the vertices of a cube forming six rings with the consequent appearance of one-ring critical point at the centre of each face of the cube. These six ring surfaces completely enclose the volume of the cube and, as a result, a cage critical point forms in the center of the cube. In Fig. 1.2a, the reader may also note the marked curvature of the bond paths in cubane, indicative of a significant ring strain in this unstable molecule.

All cage critical points reported in the literature until 2005 were found to be enclosed by at least three ring surfaces, as stated by Bader in 1990 [1]: "*While it is mathematically possible for a cage to be bounded by only two ring surfaces, the minimum number found in an actual molecule so far is three, as in bicyclo [1.1.1] pentane, for example*", a statement reiterated in 2000 [3]. In Fig. 1.2b there is nothing unusual about the aromatic system, but the nuclei of the two fluorine atoms in the "Fjord region" are linked by a bond path [4] closing a seven-membered ring which has quite an unusual topology – it gives rise to *two ring* critical points and *a cage* critical point [5]. We have, thus, recently reported the first example of an actual molecular system in which a cage is bounded by only *two* ring surfaces [5]. Such a CCP (enclosed by two ring surfaces) arises in all the studied derivatives of 1,12-difluoro[4]helicenes [5]. In these molecules, the seven-membered ring in the Fjord region is so distorted out of planarity that its ring surface splits into two, giving rise to this CCP [5]. In all cases, the Poincaré–Hopf relationship is satisfied [5].

1.3 The Topology of the Electron Density Dictates the Form of Atoms in Molecules

The pronounced maxima in the electron density at the positions of the nuclei give rise to a rich topology. This topology embodies a natural partitioning of the molecular space into separate mononuclear regions, Ω, identified as atoms in molecules. The surface bounding an atom in a molecule is one of zero flux in the gradient vector field of the electron density, i.e. it is not crossed by any of the gradient vectors $[\nabla\rho(\mathbf{r})]$ at any point, a statement which is equivalent to satisfying the condition:

$$\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0, \quad \text{for all } \mathbf{r} \text{ belonging to the surface } S(\Omega) \tag{6}$$

where $\mathbf{r}$ is the position vector and $\mathbf{n}(\mathbf{r})$ the unit vector normal to the surface $S(\Omega)$.

The plot in Fig. 1.3a represents the electron density and its gradient vector field in the molecular plane of BF_3. The figure contrasts the zero-flux surfaces which partition the molecular space into separate mononuclear "atomic basins" and an arbitrary surface cutting though the density. The left side of Fig. 1.3a is a contour

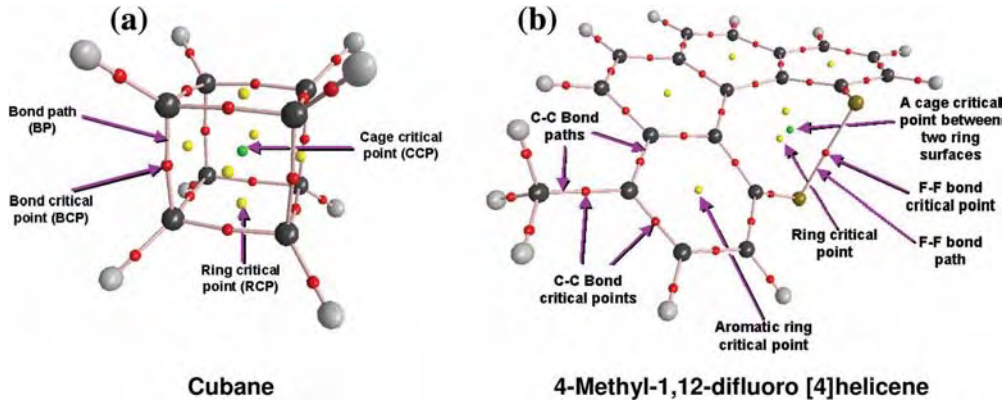


Fig. 1.2 The molecular graph of (a) cubane and (b) 4-methyl-1,12-difluoro[4]helicene showing the bond paths (lines) and the different critical points: nuclear (color-coded by element: C = black, H = grey, F = golden), bond (small red dots), ring (yellow dots), and cage (green dots) critical points.

plot of $\rho(\mathbf{r})$, the contours decreasing in value from the nuclei outward. Instead of plotting $\rho(\mathbf{r})$ in the right half of Fig. 1.3a (which is a mirror image of the left side by virtue of the molecular symmetry), we have depicted, instead, the corresponding gradient vector field $\nabla\rho(\mathbf{r})$. The gradient vector field lines partition the molecular space naturally into three fluorine basins and a central boron basin (Fig. 1.3a).

Gradient vector field lines belonging to an atomic basin all converge to *one* nucleus which acts as an attractor to these gradient vector field lines. In doing so, these gradient vector field lines sweep a portion of physical space associated with one nucleus and which is identified as the basin of an atom in a molecule (AIM). Three-dimensional volume renderings of the atoms and groups of atoms within the BF_3 molecule are shown in Fig. 1.3b. An atom in a molecule is defined as the union of a nucleus and its associated basin. Each basin is bounded by one (or by the union of a number of) zero-flux surface(s) one of which may occur at infinity. An atom in a molecule may be defined, alternatively and equivalently, as a region of space bounded by one or more zero-flux surface(s).

Occasionally, local maxima in the electron density can occur at positions other than those of atomic nuclei, especially in metals [6, 7] and semiconductors [8, 9], but also in systems such as the solvated electron [10] and at the positions of defects in crystals and color F-centers [11]. The non-nuclear maxima, also known as non-nuclear attractors (NNA), are topologically indistinguishable from the nuclear maxima. Just like a nucleus, an NNA is associated with a basin swept by gradient vector field lines and is bounded by a zero-flux surface. Consequently, NNA

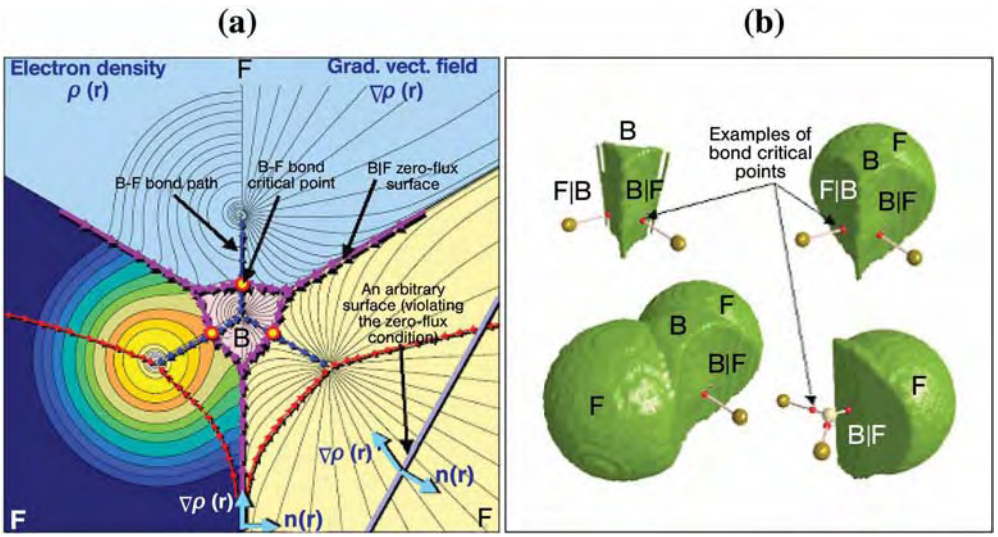


Fig. 1.3 (a) The electron density (left) and the gradient vector field (right) of the density in the molecular plane of BF_3. The blue arrows connecting the nuclei trace the bond paths. The magenta arrows delimiting atomic basins trace the intersections of the zero-flux surfaces with the plane. The contours increase from the outermost 0.001 au contour followed by 2×10^n, 4×10^n, and 8×10^n au with n starting at -3 and increasing in steps of unity. The small circles drawn on the three bond paths are the B–F bond critical points (BCP). The intersection of an arbitrary surface with the plane of the figure, the straight line on the lower right part of (a), is shown to be crossed by gradient vectors and is contrasted with a zero-flux surface. (b) Four three-dimensional renderings of the density of atoms and groupings of atoms in BF_3. The outer surface is the 0.002 au isodensity envelope. The zero-flux surfaces are denoted by the vertical bars between the atomic symbols. Large spheres represent the nuclei of the fluorine atoms (golden) and of the boron atom (blue–gray). The lines linking the nuclei represent the bond paths. The BCPs are denoted by the small red dots. A BCP always lies on the zero-flux surface shared by the two bonded atoms.

basins constitute proper open quantum systems and are therefore termed ''pseudo-atoms''. Pseudo-atoms can be bonded (i.e. share a common interatomic zero-flux surface, a bond critical point, and a bond path) to atoms and other pseudo-atoms in a molecule. Non-nuclear attractors and their basins are of great importance in characterizing metallic bonding and are of substantial theoretical interest. A detailed discussion of NNA can be found in Chapter 7 of this book.

There is a unique set of gradient vectors lines which originate at infinity and terminate at a point *between* two bonded atoms, the lines of this set fall by definition on the zero-flux surface because they satisfy Eq. (6) locally. It should be noted that the three zero-flux surfaces depicted in Fig. 1.3 are between the boron and fluorine atoms, the boron atom being bounded three zero-flux surfaces which merge in pairs at infinity between fluorine basins. There are no zero-flux surfaces

between any pair of fluorine atoms in BF_3, these surfaces only exist between bonded atoms and are characteristic of bonding interactions.

The topological definition of an atom follows from the boundary condition expressed in Eq. (6) and which applies to every point on the surface. This real space partitioning of the electron density has been shown to be rooted in quantum mechanics bringing into coincidence the topological definition of an atom in a molecule with that of a proper open quantum system (see Chapter 2 and also the detailed derivation of the quantum mechanics of proper open systems [12] from Schwinger's principle of stationary action [13]).

1.4
The Bond and Virial Paths, and the Molecular and Virial Graphs

The presence of an interatomic zero-flux surface between any two bonded atoms in a molecule is always accompanied by another key topological feature – there is, *in real space*, a single line of *locally* maximum density, termed the "bond path" (BP), linking their nuclei. The bond path is a universal indicator of chemical bonding of all kinds; weak, strong, closed-shell, and open-shell interactions [14]. The point on the bond path with the lowest value of the electron density (minimum along the path) is the bond critical point (BCP) and it is at that point where the bond path intersects the zero-flux surface separating the two bonded atoms.

The collection of bond paths linking the nuclei of bonded atoms in an equilibrium geometry, with the associated critical points, is known as the *molecular graph*. (In a non-equilibrium geometry, lines of maximum electron density linking the nuclei are known as "atomic interaction lines", because these may or may not persist when the geometry is energy-minimized, i.e. optimized.) The molecular graph provides an unambiguous definition of the "molecular structure" and can thus be used to locate changes in structure along a reaction path.

Mirroring every molecular graph is a "shadow" graph, again in real space, but this time the graph is defined by a set of lines of maximally negative *potential energy density*. In other words, there is a single line of maximally negative potential energy density linking the same attractors which share a bond path [15]. This line of "maximum stability" in real space is termed a "*virial path*". The collection of virial paths and the associated critical points constitute the *virial graph*. The virial graph defines the same molecular structure as the molecular graph, the virial field and the electron density being homeomorphic [15].

Figure 1.4 shows the chemical structure and the molecular and virial graphs of the phenanthrene molecule. This polycyclic aromatic hydrocarbon molecule has a bond path between the two hydrogen atoms in the bay region, a mode of closed-shell bonding which has been recently characterized in detail and termed hydrogen–hydrogen bonding (to be contrasted with dihydrogen bonding) [16, 17]. The virial graph is shown to faithfully map each bond path with a corresponding virial path including the bond path of a weak closed-shell bonding interaction such as the hydrogen–hydrogen bonding interaction (Fig. 1.4).

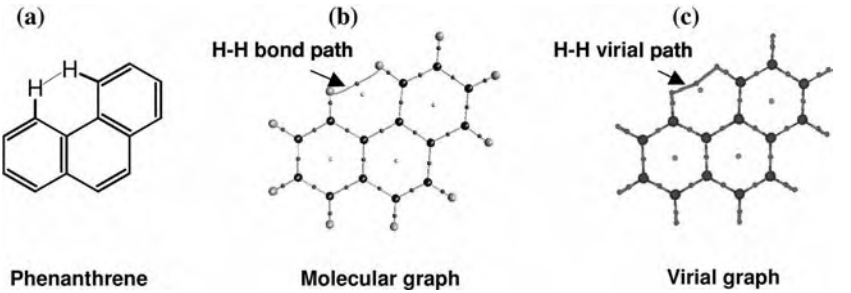


Fig. 1.4 (a) The chemical structure of phenanthrene. (b) The molecular graph of phenathrene showing the collection of bond paths and associated critical points. (c) The corresponding virial graph.

We conclude this section by stating that *atoms that are chemically bonded have their nuclei linked by a (single) bond path and by an accompanying virial path and they share a bond critical point and a common interatomic zero-flux surface.*

1.5 The Atomic Partitioning of Molecular Properties

The quantum theory of atoms in molecules is a generalization of quantum mechanics to open quantum systems. Bader has shown that the topological partitioning of the molecules into atomic basins is essential for development of the quantum mechanics of open systems [12]. The zero-flux condition, Eq. (6), is the necessary constraint for the application of Schwinger's principle of stationary action [13] to part of a quantum system [12].

The partitioning of the molecular space into atomic basins enables the partitioning of electronic properties into atomic contributions in one consistent theoretical framework. Among the properties often discussed are the atomic charges and higher multipolar electric polarizations, atomic volumes, atomic total energies (and the different contributions to the atomic energies), and the electron localization within one basin or delocalization between two basins [1, 18].

The expectation value of an operator averaged over all space is given by the sum of the expectation values of this operator averaged over all the atoms in the molecule or the crystal, in atomic units:

$$\langle \hat{O} \rangle_{molecule} = \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} \left(N \int_{\Omega_i} \left\{ \int \frac{1}{2} [\Psi^* \hat{O} \Psi + (\hat{O}\Psi)^* \Psi] \, d\tau' \right\} d\mathbf{r} \right) \tag{7a}$$

$$= \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} \left(\int_{\Omega_i} \rho_O \, d\mathbf{r} \right) = \sum_{i}^{\substack{all\ atoms\\ in\ the\\ molecule}} O(\Omega_i) \tag{7b}$$

where $\langle \hat{O} \rangle_{molecule}$ is the molecular expectation value of the operator $\hat{O}$, $O(\Omega_i)$ is the average of this operator over an atom Ω_i, and where the sum runs over all the atoms in the molecule or crystal. Integration over the coordinates of all electrons but one and summation over all spins is symbolized by $\int d\tau'$. Equation (7b) implies that any molecular property O which can be expressed in terms of a corresponding property density in space $\rho_O(\mathbf{r})$ can be written as a sum of atomic contributions obtained by averaging the appropriate operator over the volume of the atom, i.e. it exhibits atomic additivity.

1.6
The Nodal Surface in the Laplacian as the Reactive Surface of a Molecule

Because the Laplacian is essentially a second derivative, its sign indicates regions of local electronic charge concentration or depletion with respect to the immediate neighborhood. Thus, where $\nabla^2\rho(\mathbf{r}) > 0$ the density is locally depleted and expanded relative to its average distribution; where $\nabla^2\rho(\mathbf{r}) < 0$ the density is locally concentrated, tightly bound, and compressed relative to its average distribution. A local charge concentration behaves as a Lewis base (electron donor) whereas a local charge depletion acts as a Lewis acid (electron acceptor).

The Laplacian reproduces the spherical shell structure of isolated atoms in terms of alternating shells of charge concentration followed by shells of charge depletion [19, 20]. The spherical nodes in the Laplacian are envelopes bounding regions of density depletion or concentration. The outer shell of charge concentration, which is followed by a shell of charge depletion extending to infinity, is called the valence shell charge concentration (VSCC). When an atom is involved in bonding the spherical symmetry of the VSCC is broken. A chemical reaction corresponds to the combination of a "lump" in the VSCC of the base with a "hole" in the VSCC of the acid.

Covalently bonded atoms have bonding charge concentrated in the region between their nuclei. In addition to bonding charge concentrations, lone pairs are associated with non-bonding charge concentrations. These observations reflect an underlying mapping between the Laplacian of the electron density and the Laplacian of the conditional pair density when electrons tend to be localized [21].

The Laplacian of the density is characterized by a rich topology which provides a basis for the VSEPR model [22–24] of molecular geometry [1, 25–27]. More details on this topic are available elsewhere [1, 3, 25–29] and Chapter 19 explores the use of the reactive surface in drug design and drug–receptor molecular complementarity.

1.7
Bond Properties

A zero-flux surface is defined by a particular set of $\nabla\rho(\mathbf{r})$ trajectories all the members of which terminate at a single point, the bond critical point, where

$\nabla\rho(\mathbf{r}) = 0$. There is one BCP between each pair of atoms that are bonded, i.e., two atoms linked by a bond path and sharing a common interatomic zero-flux surface. In addition to the set of trajectories which terminate at the BCP and define an interatomic surface, a pair of trajectories originates at the BCP with each member of the pair terminating at one of the nuclei of the chemically bonded atoms. This latter pair of trajectories defines the bond path [14]. Chemical bonding interactions are characterized and classified according to the properties of the electron and energy densities at the BCP, collectively know as ''bond properties''.

1.7.1
The Electron Density at the BCP (ρ_b)

The strength of a chemical bond, its bond order (BO), is reflected in the electron density at the BCP (ρ_b) [1]:

$$\mathrm{BO} = \exp[A(\rho_b - B)] \tag{8}$$

where A and B are constants which depend on the nature of the bonded atoms. In general, ρ_b is greater than 0.20 au in shared (covalent) bonding and less than 0.10 au in a closed-shell interaction (for example ionic, van der Waals, hydrogen, dihydrogen, H–H bonding, etc.). ρ_b has been shown to be strongly correlated with the binding energy for several types of bonding interaction [30–36] and with the bond length of S–S bonding interactions [37]. Proposals to generalize Eq. (8) by including more than two elements in the same fitting have recently appeared in the literature [38, 39].

1.7.2
The Bonded Radius of an Atom (r_b), and the Bond Path Length

The distance of a BCP from nucleus A determines the ''bonded radius'' of atom A relative to the interaction defined by the BCP, and is denoted r_b(A). If the bond path is coincident with the internuclear axis, then the sum of the two associated bond radii, termed the bond path length, equals the bond length. If, however, the bond path is curved, or strained chemically, the bond path length will exceed the bond length. Examples of this latter behavior are found for hydrogen-bonded interactions and for bonding within strained cyclic molecules (e.g. the curved C–C bond paths in the cubane molecule, Fig. 1.2a).

1.7.3
The Laplacian of the Electron Density at the BCP ($\nabla^2\rho_b$)

The Laplacian at the BCP is the sum of the three curvatures of the density at the critical point (Eq. 4), the two perpendicular to the bond path, λ_1 and λ_2, being negative (by convention, $|\lambda_1| > |\lambda_2|$) whereas the third, λ_3, lying along the bond

path, is positive. The negative curvatures measure the extent to which the density is concentrated along the bond path and the positive curvature measures the extent to which it is depleted in the region of the interatomic surface and concentrated in the individual atomic basins.

In covalent bonding the two negative curvatures are dominant and $\nabla^2\rho_b < 0$, for example, $\nabla^2\rho_b = -1.1$ au for a typical C–H bond. In contrast, in closed-shell bonding, for example ionic, hydrogen-bonding or van der Walls interactions, the interaction is characterized by a depletion of density in the region of contact of the two atoms and $\nabla^2\rho_b > 0$. An N–(H···O)=C hydrogen bond, for instance, is characterized by $\nabla^2\rho_b = +0.03$ au. In strongly polar bonding, (e.g. C–X, where X = O, N, F), there is a significant accumulation of electron density between the nuclei, as in all shared interactions, but the Laplacian in this type of bonding can be of either sign.

1.7.4
The Bond Ellipticity (ε)

The ellipticity measures the extent to which density is preferentially accumulated in a given plane containing the bond path. The ellipticity is defined as:

$$\varepsilon = \frac{\lambda_1}{\lambda_2} - 1 \quad (\text{where } |\lambda_1| \geq |\lambda_2|) \tag{9}$$

If $\lambda_1 = \lambda_2$, then $\varepsilon = 0$, and the bond is cylindrically symmetrical; examples are the C–C single bond in ethane or the triple bond in acetylene. Thus, ε is a measure of the π-character of the bonding up to the limit of the "double bond" for which the ellipticity reaches a maximum. On going from a double to a triple bond, the trend is reversed and the ellipticity decreases with increasing bond order, because at the limit of BO = 3 the bonding regains its cylindrical symmetry (two π-bonding interactions in two orthogonal planes in addition to a cylindrically symmetric σ-bonding interaction). The ellipticity of an aromatic bond is ca. 0.23 in benzene and that of a formal double bond is ca. 0.45 in ethylene.

1.7.5
Energy Densities at the BCP

Energy densities require information contained in the one-electron density matrix (and not just the density, its diagonal elements). The energy densities (potential, kinetic, and total) are used to summarize the mechanics of a bonding interaction.

The potential energy density, $\mathcal{V}(\mathbf{r})$, also known as the virial field, is the average effective potential field experienced by a single electron at point $\mathbf{r}$ in a many-particle system. The virial field evaluated at any point in space is always negative and its integral over all space yields the total potential energy of the molecule. The local statement of the virial theorem expresses the relationship between the

virial field, the kinetic energy density, and the Laplacian, which when written for a stationary state is [1, 12, 40]:

$$\left(\frac{\hbar^2}{4m}\right)\nabla^2\rho(\mathbf{r}) = 2G(\mathbf{r}) + \mathscr{V}(\mathbf{r}) \tag{10}$$

where

$$G(\mathbf{r}) = \frac{\hbar^2}{2m} N \int d\tau' \nabla\Psi^* \cdot \nabla\Psi \tag{11}$$

and where $G(\mathbf{r})$ is the gradient kinetic energy density and Ψ is an antisymmetric many-electron wavefunction.

Because we always have $G(\mathbf{r}) > 0$ and $\mathscr{V}(\mathbf{r}) < 0$, the local virial theorem when applied at a BCP implies that interactions for which $\nabla^2\rho_b < 0$ are dominated by a local reduction of the potential energy. Conversely, interactions for which $\nabla^2\rho_b > 0$ are dominated by a local excess in the kinetic energy.

To compare the kinetic and potential energy densities on an equal footing (instead of the 2:1 virial ratio) Cremer and Kraka [41] proposed evaluating the total electronic energy density $[H(\mathbf{r}) = G(\mathbf{r}) + \mathscr{V}(\mathbf{r})]$ at the BCP:

$$H_b = G_b + \mathscr{V}_b \tag{12}$$

The total energy density yields the total electronic energy when integrated over all space. H_b is negative for interactions with significant sharing of electrons, its magnitude reflecting the "covalence" of the interaction [41].

1.7.6
Electron Delocalization between Bonded Atoms: A Direct Measure of Bond Order

The number of electron pairs *shared* between two bonded atoms is often called the *bond order*. QTAIM provides a bookkeeping of the number of pairs shared between two atoms by integrating the exchange density once over each of the two atomic basins. This property may as well be classified under "atomic properties" because it involves the double integration of the exchange density over the basins of two atoms, but, because it "counts" the number of electron pairs shared between two atoms, when reported for *bonded* atoms, it can be regarded as a bond property.

The magnitude of the exchange of the electrons in the basin of atom A with those in the basin of atom B is termed the delocalization index between them, $\delta(A, B)$, and is defined for a closed-shell system as [42]:

$$\delta(A, B) = 2|F^{\alpha}(A, B)| + 2|F^{\beta}(A, B)| \tag{13}$$

where the Fermi correlation is defined as:

$$F^{\sigma}(A,B) = -\sum_{i}\sum_{j}\int_{A} d\mathbf{r}_1 \int_{B} d\mathbf{r}_2 \{\phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_i(\mathbf{r}_2)\}$$
$$= -\sum_{i}\sum_{j} S_{ij}(A)S_{ji}(B) \tag{14}$$

where $S_{ij}(\Omega) = S_{ji}(\Omega)$ is the overlap integral of two spin orbitals over a region Ω and σ represents spin (α or β).

The second-order density matrix obtained from a configuration interaction (CI) calculation can also be expressed in terms of products of basis functions multiplied by the appropriate coefficients enabling one to express the integrated pair density in terms of overlap contributions. Thus, terms similar to those in Eq. (14) multiplied by the appropriate coefficients appear in the CI expression for $F^{\sigma}(A,B)$ and electron delocalization is still described in terms of the exchange of electrons between molecular orbitals, but this time in a wavefunction incorporating Coulomb in addition to Fermi correlation [42].

If the double integration in Eq. (14) is performed over only one atomic basin, say atom A, this would yield the total Fermi correlation for the electrons in region A [43]:

$$F^{\sigma}(A,A) = \int_{A} d\mathbf{r}_1 \int_{A} d\mathbf{r}_2 \rho^{\sigma}(\mathbf{r}_1)h^{\sigma}(\mathbf{r}_1,\mathbf{r}_2) \tag{15}$$

where its limiting value is $-N^{\sigma}(A)$, the negative of the σ-spin population of atom A, i.e. the number of σ electrons in A being totally localized within this atom because all remaining σ-spin density would then be excluded from A. In other words, if this limiting value is reached, it implies that the electrons in A do not exchange with electrons outside A. Thus a *localization index* [$\lambda(A)$] is defined as:

$$\lambda(A,A) = |F^{\alpha}(A,A)| + |F^{\beta}(A,A)| \tag{16}$$

The limit of total localization, while approached quite closely ($\geq$95%) in ionic systems, cannot usually be achieved and one finds that $|F^{\sigma}(A,A)| < N^{\sigma}(A)$, indicating that the electrons in region A always exchange, to some extent, with electrons outside the boundaries of A, i.e., they are delocalized.

Because the Fermi correlation counts all electrons, the sum of the localization indices and half of all the delocalization indices is N, the total number of electrons in the molecule. This, in turn, provides a measure of how these electrons are localized within the individual atomic basins and delocalized between them, in effect resulting in bookkeeping of electrons in the molecule:

$$N(A) = \lambda(A) + \frac{1}{2}\sum_{B\neq A}\delta(A,B) \tag{17}$$

How closely the sum of the localization and the delocalization indices (Eq. 17) recovers the total molecular electron population is a global measure of the quality of the atomic integrations.

The localization and delocalization indices can be calculated from the atomic overlap matrices using readily available software such as AIMDELOC [44] or LI-DICALC [45, 46].

It is important to realize that a delocalization index can be calculated between *any* pair of atoms whether bonded or not. When $\delta(A, B)$ is calculated between *bonded* atoms it yields a measure of the bond order between them if the electron pairs are equally shared (i.e. there is no appreciable charge transfer) [42, 47].

Because ρ_b and the bond order are strongly correlated (Eq. 8), Matta and Hernández-Trujillo [48] suggested calibrating this correlation using the delocalization index rather than arbitrarily assigned bond orders:

$$\delta(A, B) = \exp[A(\rho_b - B)] \tag{18}$$

Equation (18) enables calibration of experimental ρ_b with delocalization indices obtained by calculation. The fitted equation can then be used to obtain experimental estimates for information on electron sharing contained in a full density matrix, information which is not accessible in a conventional X-ray diffraction experiment, from experimentally determined ρ_b [48]. Data for the 21 carbon–carbon bonds in the estrone hormone could be fitted to the following equation [49]:

$$\delta(\mathrm{C}, \mathrm{C}') = \exp\{4.7427 \times [\rho_b(\text{in a.u.}) - 0.2538]\} \tag{19}$$

with $r^2 = 0.939$, a variance of 0.002, and a root mean square deviation of 0.010, and in which $\delta(\mathrm{C}, \mathrm{C}')$ were calculated at the B3LYP/6-311++$G(d, p)$ level and ρ_b are the experimentally determined electron density values at the C–C BCPs.

1.8 Atomic Properties

The average of a property O over an atomic basin Ω, $O(\Omega)$, is calculated from:

$$O(\Omega) = \langle \hat{O} \rangle_\Omega = \frac{N}{2} \int_\Omega d\mathbf{r} \int d\tau' [\Psi^* \hat{O} \Psi + (\hat{O}\Psi)^* \Psi] \tag{20}$$

where $\hat{O}$ is a one-electron operator or a sum of one-electron operators. Some examples of commonly computed atomic properties are discussed in the subsections below.

1.8.1
Atomic Electron Population [$N(\Omega)$] and Charge [$q(\Omega)$]

The total electron population of an atom in a molecule is obtained by setting $\hat{O} = \hat{1}$ in Eq. (20). This yields:

$$N(\Omega) = \int_{\Omega} \rho(\mathbf{r})\, d\mathbf{r} \tag{21}$$

which can also be expressed explicitly in terms of the separate spin populations as the expectation value of the number operator, an integral operator, averaged over a proper open quantum subsystem:

$$N(\Omega) = \sum_{i} [\langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\alpha} + \langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\beta}] \tag{22}$$

in which the separate spin populations are given by:

$$\langle \psi_i(\mathbf{r}) | \psi_i(\mathbf{r}) \rangle_{\Omega}^{\sigma} = \int_{\Omega} \psi_i^{\sigma *}(\mathbf{r}) \psi_i^{\sigma}(\mathbf{r})\, d\mathbf{r} \equiv S_{ii}^{\sigma}(\Omega) \tag{23}$$

where σ refers to either α-spin or β-spin, and $S_{ii}^{\sigma}(\Omega)$ is the ith diagonal element of the atomic overlap matrix.

The atomic charge is obtained by subtracting $N(\Omega)$ from the nuclear charge Z_{Ω}:

$$q(\Omega) = Z_{\Omega} - N(\Omega) \tag{24}$$

Because of the manner by which atomic populations are defined, Eqs (22) and (23), QTAIM populations and charges are true quantum expectation values. That is, they are "observables" in the quantum mechanical sense [18, 50]. Observables are not necessarily measurable in practice, but any measurable quantity is an observable or can be expressed in terms of one or more observables. Indirect experimental evidence lends strong support to the physical nature of QTAIM atomic populations and charges [51] (see also Section 1.9.2).

The deviation of the sum of the atomic populations (or charges) from the corresponding molecular value is an indicator of the quality of the numerical integrations. Deviations of less than ca. 0.001–0.002 electrons are regarded as acceptable for molecules of medium size (up to ~100 first to third row atoms).

1.8.2
Atomic Volume [Vol.(Ω)]

The atomic volume is defined as the space bounded by the intersection of the zero-flux surface(s) bounding the atom from the molecular interior and a chosen

outer isodensity envelope (if a side of this atom's basin extends to infinity). While a molecule extends in principle to infinity, an outer isodensity of $\rho(\mathbf{r}) = 0.001$ au is usually chosen as its outer bounding surface for two reasons:

1. this isosurface closely recovers the experimental van der Waals volumes in the gas phase, and
2. it usually encloses more than 99% of the electron population of the molecule [1].

The van der Waals surface in condensed phases is closer to the 0.002 au isodensity envelope [1].

1.8.3 Kinetic Energy [$T(\Omega)$]

There are at least two forms of the kinetic energy operator [52] with two corresponding expressions for the atomic average of the kinetic energy, the Schrödinger kinetic energy:

$$K(\Omega) = -\frac{\hbar^2}{4m} N \int_\Omega d\mathbf{r} \int d\tau' [\Psi \nabla^2 \Psi^* + \Psi^* \nabla^2 \Psi] \tag{25}$$

and the gradient kinetic energy:

$$G(\Omega) = \frac{\hbar^2}{2m} N \int_\Omega d\mathbf{r} \int d\tau' \nabla_i \Psi^* \cdot \nabla_i \Psi \tag{26}$$

For the total system and for a proper open quantum system, Eqs (25) and (26) must yield an identical value for the kinetic energy, of course, i.e. $K(\Omega) = G(\Omega) = T(\Omega)$. Because the difference between $K(\Omega)$ and $G(\Omega)$ should vanish for an atom in a molecule, the (small) departure from zero of this difference as gauged by the Laplacian (Section 1.8.4) is a measure of the numerical accuracy of the atomic integrations.

1.8.4 Laplacian [$L(\Omega)$]

The Laplacian function has the dimensions of electrons $\times$ (length)$^{-5}$. Because of the zero-flux boundary condition, Eq. (6), the Laplacian of the electron density, vanishes when integrated over an atomic basin, as can be seen from:

$$\begin{aligned} L(\Omega) &= K(\Omega) - G(\Omega) \\ &= -\frac{\hbar^2}{4m} \int_\Omega d\mathbf{r} [\nabla^2 \rho(\mathbf{r})] \\ &= -\frac{\hbar^2}{4m} \int dS(\Omega, \mathbf{r}) \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \end{aligned} \tag{27}$$

the last equality is valid only for the total system or if the integration is performed over a proper open quantum system bounded by zero-flux surfaces.

How close the integrated Laplacian approaches zero is often used as an indicator of the numerical accuracy of atomic integrations. Deviations from zero are a measure of integration error. $L(\Omega) \leq$ ca. 1.0×10^{-3} au for second and third-row atoms and $L(\Omega) \leq$ ca. 1.0×10^{-4} au for hydrogen atoms are regarded as acceptable and are usually paralleled by atomic energies which add up to within a kcal mol^{-1} of the directly calculated molecular total energy for a medium size molecule ($\sim$100 atoms or fewer). The smaller $L(\Omega)$ the better the quality of an atomic integration.

1.8.5 Total Atomic Energy [$E_e(\Omega)$]

The partitioning of the total molecular energy into a set of additive atomic energies is a non-trivial problem that was solved by Bader [1]. To see the difficulties in partitioning the total energy, one may ask, for instance, how can the nuclear–nuclear repulsion contribution to the total molecular energy be partitioned on an atom-by-atom basis?

The kinetic energy density can be expressed:

$$K(\mathbf{r}) = -\frac{\hbar^2}{4m} N \int d\tau' [\Psi \nabla^2 \Psi^* + \Psi^* \nabla^2 \Psi] \tag{28}$$

which when compared with Eq. (11) yields:

$$K(\mathbf{r}) = G(\mathbf{r}) - \frac{\hbar^2}{4m} \nabla^2 \rho(\mathbf{r}) \tag{29}$$

It is clear from Eq. (29) that the integral of the kinetic energy densities $K(\mathbf{r})$ and $G(\mathbf{r})$ over a volume ω would usually yield different values because the integral of the Laplacian does not usually vanish when integrated over an arbitrary volume, in which case the kinetic energy is *not* well defined. The kinetic energy is well defined if, and only if, the integral of the Laplacian term vanishes, i.e. when this integral is performed over the total system or over an atomic basin bounded by a zero-flux surface. Integrating Eq. (29) over ω, one obtains:

$$K(\omega) = G(\omega) - \frac{\hbar^2}{4m} N \int_\omega d\mathbf{r}\, \nabla \cdot \nabla \rho \tag{30}$$

Using Gauss's theorem, the volume integral in Eq. (30) can be transformed into a surface integral:

$$K(\omega) = G(\omega) - \frac{\hbar^2}{4m} N \int dS(\omega, \mathbf{r}) \nabla \rho \cdot \mathbf{n}(\mathbf{r}) \tag{31}$$

Now it is clear that the second term in the R.H.S. will vanish only for systems bounded by a zero-flux surface satisfying Eq. (6) (or for the whole system, because the Laplacian integrated over the entire space also vanishes). Thus only the total system and proper open sub-systems will have a definite kinetic energy. A proper open system (one bounded by a zero-flux surface and/or infinity) will be referred to as Ω to distinguish it from an arbitrary bounded region of space ω. For such a proper open system one has:

$$K(\Omega) = G(\Omega) = T(\Omega) \tag{32}$$

and, because the integral of the Laplacian vanishes over Ω, the integral of the local statement of the virial theorem (Eq. 10) over Ω yields the *atomic virial theorem*:

$$-2T(\Omega) = \mathscr{V}(\Omega) \tag{33}$$

where the $\mathscr{V}(\Omega)$ is the total atomic virial.

The atomic electronic energy $E_e(\Omega)$ is given by:

$$E_e(\Omega) = T(\Omega) + \mathscr{V}(\Omega) \tag{34}$$

For systems in equilibrium there are no Hellmann–Feynman forces acting on the nuclei and the virial equals the average potential energy of the molecule, i.e. $\mathscr{V} = V$. Under this condition Eq. (33) becomes:

$$-2T(\Omega) = V(\Omega) \tag{35}$$

where $V(\Omega)$ is the potential energy of atom Ω, and Eq. (34) becomes:

$$E(\Omega) = E_e(\Omega) = T(\Omega) + V(\Omega) = -T(\Omega) = \frac{1}{2}V(\Omega) \tag{36}$$

where $E(\Omega)$ is the total energy of atom Ω.

Thus, the energy of an atom in a molecule at its equilibrium geometry is obtained from the atomic statement of the virial theorem, and $E(\Omega) = -T(\Omega)$. The sum of atomic energies yields, naturally, the total energy of the molecule (obtained directly from the electronic structure calculation) to within a small numerical integration error. This additivity of the atomic energies is expressed as:

$$E_{total} = \sum_{\Omega} E(\Omega) \tag{37}$$

The result shown in Eq. (37) is remarkable. The equation expresses the partitioning of the total molecular energy into atomic contributions, a partitioning which includes, for example, the nuclear–nuclear repulsion contribution to the

molecular energy. Such partitioning of the total energy is indispensable if one is to understand the atomic origins of the energy difference between two isomers [4, 16, 53], for example, or the atomic origins of potential energy barriers [17, 54].

The deviation of the sum of the atomic energies from the directly calculated total molecular energy is another global measure of the quality of atomic integrations. A deviation of no more than ca. 1 kcal mol^{-1} is usually regarded as an indicator of accurate integrations.

The discussion above is based upon the assumption that the calculated molecular wavefunction satisfies the virial theorem exactly, i.e. the molecular virial ratio $-V/T = 2$ to infinite accuracy. In practice, the calculated virial ratio deviates slightly from this ideal value of 2 because of the truncation of the basis set, residual forces on the nuclei, and the finite nature of the convergence thresholds in a typical calculation. The manner by which AIMPAC corrects for this deviation is described in the Appendix.

1.8.6
Atomic Dipolar Polarization [$\boldsymbol{\mu}(\Omega)$]

Also known as the *first atomic electrostatic moment*, atomic dipolar polarization is the atomic space average of the electronic position vector. It is a three-dimensional vector with components and magnitude defined in Eqs (38) and (39), respectively:

$$\boldsymbol{\mu}(\Omega) = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix} = \begin{pmatrix} -e\int_\Omega x\rho(\mathbf{r})\,d\mathbf{r} \\ -e\int_\Omega y\rho(\mathbf{r})\,d\mathbf{r} \\ -e\int_\Omega z\rho(\mathbf{r})\,d\mathbf{r} \end{pmatrix} \equiv -e\int_\Omega \mathbf{r}_\Omega\rho(\mathbf{r})\,d\mathbf{r} \tag{38}$$

$$|\boldsymbol{\mu}(\Omega)| = \sqrt{{\mu_x}^2 + {\mu_y}^2 + {\mu_z}^2} \tag{39}$$

with the origin for the vector $\mathbf{r}_\Omega$ at the nucleus of atom Ω, i.e. $\mathbf{r}_\Omega = \mathbf{r} - \mathbf{R}_\Omega$, $\mathbf{r}$ being the electronic coordinates and $\mathbf{R}_\Omega$ the nuclear coordinates of atom Ω. The first moment measures the polarization of the charge density, that is to say the departure from sphericity of the electron density.

The dipolar polarization can be used to understand the origins of permanent and induced molecular dipole moments and dielectric polarization in materials [55, 56]. Carbon monoxide is an instructive example [57]. The dipole moment of the CO molecule has the (unexpected) polarity $^{\delta-}C{=}O^{\delta+}$, opposite to intuition based on the relative electronegativities of carbon and oxygen. This observation is readily explained when one considers both atomic charges and atomic dipoles. Calculated atomic charges are indeed in accordance with the expected relative electronegativities of these two atoms (an electronegativity of 2.5 for carbon and

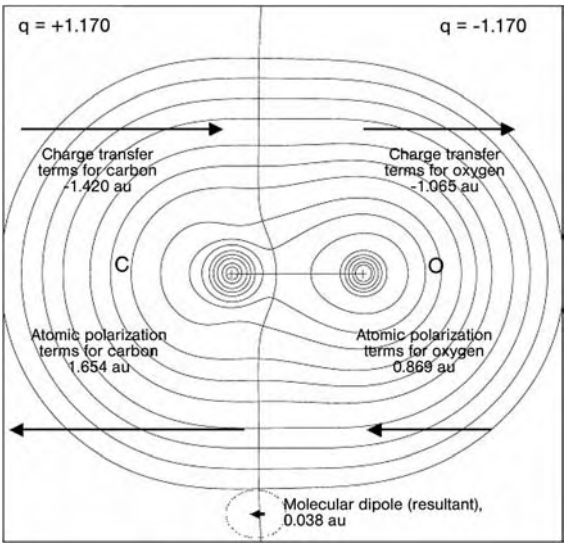

q = +1.170
q = -1.170
Charge transfer terms for carbon -1.420 au
Charge transfer terms for oxygen -1.065 au
C
O
Atomic polarization terms for carbon 1.654 au
Atomic polarization terms for oxygen 0.869 au
Molecular dipole (resultant), 0.038 au

Fig. 1.5 Contour plot of the electron density of CO, showing the magnitudes and directions of atomic and charge-transfer dipoles (arrow lengths are proportional to the dipoles magnitudes). The head of an arrow points to the negative end of a dipole. The molecular dipole moment is given by the vector sum of the charge-transfer terms ($\boldsymbol{\mu}_{CT}$) and the atomic polarizations ($\boldsymbol{\mu}_{AP}$). The directly calculated SCF molecular dipole is 0.096 debyes (D) at the B3LYP/6-311+G(3df) and the corresponding dipole obtained from group contributions is 0.096 D, (experimental: 0.110 D). (Reproduced from Ref. [57] with the permission from the *American Chemical Society*).

3.5 for oxygen) [58] with the carbon bearing a positive charge [$q(C) = +1.17$ au] and the oxygen a negative charge, resulting in a *charge-transfer dipole* with the direction $^{+1.17}\overrightarrow{C{=}O}^{-1.17}$. The electron density of each of the two atomic basins responds to this charge-transfer dipole with an opposing dipolar polarization $\overleftarrow{C}{=}\overleftarrow{O}$ with a magnitude which is not only sufficient to cancel the charge-transfer dipole but to slightly exceed it. The net result is a small dipole in complete accordance with the unexpected experimental result (Fig. 1.5). Thus, it is necessary to take the vectorial sum of both the charge transfer and the atomic polarization dipoles in defining atomic or group contributions to the molecular dipole moment [55, 56].

A program, FRAGDIP [59], is available for calculation of additive atomic and group contributions to the molecular dipole moment. As an illustration, Table 1.1 lists the group contributions to the dipole moments of several naturally occurring amino acids with their vector sum and compares this sum with the dipole moments calculated directly from an SCF calculation (second line). Each amino acid in its neutral form, with general formula R–CH(NH_2)COOH, was regarded as consisting of two groups – the side-chain (R–) and the "main chain" (–CH(NH_2)COOH). The reader can see how closely the group contributions sum to the molecular dipole.

Table 1.1 Comparison of the molecular dipole moments of some amino acids obtained from group contributions with those obtained directly from SCF calculations.[a,b] The table also compares the second letter of the genetic code of each amino acid with side-chain dipole. (Adapted from Refs. [56, 60]).

Amino acid	Second base in the mRNA genetic codon[c]	Nature of the 2nd base[d] in the codon	Side-chain dipole[a,b] (au) R–				Main-chain dipole[a,b] (au) –CH(NH_2)COOH				Total molecular dipole[a,b] (au) R–CH(NH_2)COOH			
			μ_x	μ_y	μ_z	$\lvert\mu\rvert$	μ_x	μ_y	μ_z	$\lvert\mu\rvert$	μ_x	μ_y	μ_z	$\lvert\mu\rvert$
Gly	G	I	−0.006	0.068	−0.098	0.119	0.415	−0.376	0.095	0.568	0.409	−0.307	−0.002	0.512
											0.402	−0.317	−0.001	0.512
Ile	U	II	0.130	0.012	−0.074	0.150	−0.363	0.521	−0.267	0.689	−0.233	0.533	−0.341	0.674
											−0.207	0.537	−0.269	0.635
Ala	C	II	0.034	−0.108	−0.120	0.165	0.217	0.590	−0.073	0.632	0.251	0.481	−0.193	0.576
											0.252	0.480	−0.193	0.576
Val	U	II	−0.103	0.149	−0.062	0.191	0.097	−0.621	−0.266	0.683	−0.006	−0.473	−0.328	0.575
											−0.013	−0.473	−0.319	0.571
Leu	U	II	0.069	0.182	0.081	0.211	−0.195	−0.607	0.163	0.658	−0.127	−0.424	0.244	0.506
											−0.119	−0.412	0.264	0.503
Phe	U	II	0.217	−0.142	−0.118	0.285	−0.099	0.642	−0.057	0.652	0.118	0.501	−0.175	0.543
											0.077	0.495	−0.187	0.535
Tyr	A	I	−0.292	−0.401	−0.201	0.536	0.138	−0.622	−0.065	0.641	−0.154	−1.024	−0.266	1.069
											−0.137	−1.026	−0.272	1.071
Thr	C	II	0.399	−0.259	−0.548	0.726	0.114	−0.645	0.065	0.658	0.514	−0.904	−0.483	1.147
											0.515	−0.904	−0.479	1.146
Met	U	II	−0.377	−0.636	−0.327	0.809	−0.288	−0.005	−0.382	0.479	−0.665	−0.642	−0.709	1.165
											−0.661	−0.647	−0.695	1.157

Trp	G	I	0.354	0.727	−0.140	0.820	−0.084	0.498	−0.173	0.534	0.270	1.225	−0.313	1.293
											0.245	1.233	−0.315	1.296
Cys	G	I	−0.692	0.090	0.450	0.830	−0.366	0.017	−0.479	0.603	−1.058	0.107	−0.029	1.064
											−1.047	0.118	−0.033	1.055
Ser	G/C[e]	I/II[e]	−0.823	−0.091	0.174	0.847	0.390	−0.428	−0.047	0.581	−0.433	−0.519	0.127	0.688
											−0.441	−0.528	0.130	0.700
Gln	A	I	−0.206	−1.296	−1.030	1.668	−0.056	0.436	0.695	0.823	−0.261	−0.860	−0.335	0.959
											−0.255	−0.894	−0.339	0.989
His	A	I	−1.510	−0.722	0.782	1.848	−0.380	0.097	−0.487	0.625	−1.890	−0.625	0.295	2.013
											−1.860	−0.637	0.303	1.989
Asn	A	I	−0.963	1.631	−0.197	1.904	−0.019	−0.651	−0.249	0.697	−0.982	0.980	−0.446	1.457
											−0.983	0.988	−0.458	1.467

[a] The total molecular dipole of an amino acid symbolized by R–CH(NH_2)COOH is given in the last four columns. For each amino acid, the top line in the last four columns lists the dipole moment obtained from the "side-chain" (R–) and the "main chain" (–CH(NH_2)COOH) group contributions using the FRAGDIP program [55, 56, 59], and the second line lists the corresponding dipole moment calculated directly from the SCF using the Gaussian program [61].

[b] Electronic structure calculations were performed at the HF/6-311++G(d,p)//HF/6-31+G(d) level of theory.

[c] The second base in the RNA triplet genetic code of the amino acid: A = adenine, C = cytosine, G = guanine, U = uracil.

[d] The chemical nature of the second base in the triplet code: I = purine (adenine and guanine), and II = pyrimidine (cytosine, uracil).

[e] Serine is the only amino acid which has a degenerate genetic code in the second position, all other amino acids have a unique base in the second position in all of their synonymous codons.

Further, the listings in Table 1.1 have been sorted in terms of the magnitude of the side-chain dipole magnitude, a sorting that reveals a striking regularity in the genetic code. Most amino acids listed in the upper part of the table with side-chain dipole magnitude less than 0.81 au (i.e. with non-polar side-chains) are encoded by a genetic triplet code having a pyrimidine base as the middle letter in the mRNA codon (except glycine, which lacks a side-chain, and tyrosine). On the other hand, most polar amino acids (having side-chain dipole magnitudes greater than 0.81 au) are encoded by a purine base, serine being the only "degenerate" amino acid, having codons of both types [56, 60]. Whereas this regularity in the genetic code has been well known for a long time, it is given a quantitative basis derived directly from the electron density distributions of the amino acids for the first time [60].

1.8.7
Atomic Quadrupolar Polarization [Q(Ω)]

The atomic quadrupolar polarization tensor is also known as the *second atomic electrostatic moment*. It is a symmetric traceless tensor defined as:

$$\mathbf{Q}(\Omega) = \begin{pmatrix} Q_{xx} & Q_{xy} & Q_{xz} \\ Q_{yx} & Q_{yy} & Q_{yz} \\ Q_{zx} & Q_{zy} & Q_{zz} \end{pmatrix}$$

$$\equiv -\frac{e}{2} \begin{pmatrix} \int_\Omega (3x_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} & 3\int_\Omega x_\Omega y_\Omega \rho(\mathbf{r})\, d\mathbf{r} & 3\int_\Omega x_\Omega z_\Omega \rho(\mathbf{r})\, d\mathbf{r} \\ 3\int_\Omega y_\Omega x_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega (3y_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega y_\Omega z_\Omega \rho(\mathbf{r})\, d\mathbf{r} \\ 3\int_\Omega z_\Omega x_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega z_\Omega y_\Omega \rho(\mathbf{r})\, d\mathbf{r} & \int_\Omega (3z_\Omega^2 - r_\Omega)\rho(\mathbf{r})\, d\mathbf{r} \end{pmatrix} \tag{40}$$

where, as for the first moment, the origin is placed at the nucleus. If the atomic electron density has spherical symmetry, then $\int_\Omega x_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \int_\Omega y_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \int_\Omega z_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r} = \frac{1}{3}\int_\Omega \mathbf{r}_\Omega^2 \rho(\mathbf{r})\, d\mathbf{r}$, and $Q_{xx} = Q_{yy} = Q_{zz} = 0$. Thus, the quadrupole moment is another measure of the deviation of the atomic electron density from spherical symmetry. For example, if a diagonal component of $\mathbf{Q}(\Omega)$ is <0, the electron density is concentrated along that axis, and vice versa. It is always possible to find a coordinate system such that the original tensor in Eq. (40) [$\mathbf{Q}(\Omega)$] is diagonalized [$\mathcal{Q}(\Omega)$]. The diagonalization of $\mathbf{Q}(\Omega)$ corresponds to a rotation of the original coordinate system. The diagonalized quadrupole tensor corresponding to Eq. (40) is written:

$$\mathcal{Q}(\Omega) = \begin{pmatrix} \mathcal{Q}_{x'x'} & 0 & 0 \\ 0 & \mathcal{Q}_{y'y'} & 0 \\ 0 & 0 & \mathcal{Q}_{z'z'} \end{pmatrix} \tag{41}$$

where $\mathcal{Q}_{x'x'}$, $\mathcal{Q}_{y'y'}$, and $\mathcal{Q}_{z'z'}$ are the principal values of the quadrupole moment with regard to the principal (rotated) axes, the x', y', and z' axes, which correspond to axes of symmetry if they exist in the electron density distribution (the primes will be dropped for simplicity).

The traceless property of the tensor defined in Eq. (40) (or in its diagonalized form, Eq. 41) is a consequence of the equality:

$$\mathbf{r}_\Omega^2 = x_\Omega^2 + y_\Omega^2 + z_\Omega^2 \tag{42}$$

which is always true in any coordinate system. Therefore:

$$(Q_{xx} + Q_{yy} + Q_{zz}) = (\mathcal{Q}_{xx} + \mathcal{Q}_{yy} + \mathcal{Q}_{zz}) = 0 \tag{43}$$

and only five independent components completely specify $\mathbf{Q}(\Omega)$ in the original coordinate system and only two are sufficient to specify its diagonalized form $\mathcal{Q}(\Omega)$.

Finally, the magnitude of the quadrupolar polarization moment is defined as [62]:

$$|\mathbf{Q}| = \sqrt{\frac{2}{3}(Q_{xx}^2 + Q_{yy}^2 + Q_{zz}^2)} = \sqrt{\frac{2}{3}(\mathcal{Q}_{xx}^2 + \mathcal{Q}_{yy}^2 + \mathcal{Q}_{zz}^2)} \tag{44}$$

1.9
"Practical" Uses and Utility of QTAIM Bond and Atomic Properties

1.9.1
The Use of QTAIM Bond Critical Point Properties

Several QTAIM bond properties have been shown to be correlated with experimental molecular properties. For example, the electron density at the BCP, ρ_b, has been shown on several occasions to be strongly correlated with the bond energies, and hence provide a measure of bond order (Eq. 8) [1, 30]; the potential energy density at the BCP has been shown to be highly correlated with hydrogen bond energies [32]; full interaction potentials in hydrogen bonds were recovered from the potential energy density at the BCP [63]; π–π stacking interactions in benzene dimers and in stacked DNA bases and base-pairs have been found to be highly correlated to BCP and cage critical point data between π-stacked monomers [64–66].

The use of BCP properties in drug design is a field pioneered by Popelier and coworkers. These authors proposed the construction of a vector space from bond properties evaluated at the bond critical points, i.e. a point in this space is specified by a set of bond properties [67–70]. This space was used as a basis for comparing related molecules, the smaller the distance between two molecules in this

space the more they are similar. Quantification of molecular similarity in this manner has several advantages over other similarity measures (for example Carbo's similarity index [71]):

1. it is much faster because it involves no spatial integration (the density of each molecule is only sampled at the positions of the BCPs);
2. it is not dominated by nuclear maxima but rather emphasizes the more interesting chemical bonding regions of the molecule; and
3. it is not plagued with the alignment problem, in which one must often choose how to align the molecules to be compared before the integration.

The new method has been successful in accurately predicting a number of properties of several series of molecules [67–70].

1.9.2 The Use of QTAIM Atomic Properties

The review in this section follows closely Table 1 of Ref. [51].

Atomic properties have been used to recover and *directly* predict several experimentally additive atomic and group contributions to molecular properties, including, for example, heats of formation [72], magnetic susceptibility (Refs [73–77] and Chapter 3 in this book), molecular volumes [78], electric moments (Chapter 3) and polarizability [79–81], Raman intensities [79, 81–84] (see also Chapter 4), IR intensities [85–88] (see also Chapter 4), spectroscopic transition probabilities [89], dielectric polarization in crystals and molecular dipole and quadrupole moments [55, 90, 91], Wigner–Seitz cells in crystals [92], group additivity in silanes [93], and Pascal's aromatic exaltations [1]. They have also been used to provide an atomic basis for electron localization and delocalization [42, 43, 47, 94, 95].

Atomic properties have also been used *empirically* to predict several experimental properties including for example, the pK_a of weak acids from the atomic energy of the acidic hydrogen [96], a wide array of biological and physicochemical properties of the amino acids, including the genetic code itself, and the effects of mutation on protein stability [60], protein retention times [97], HPLC column capacity factors of high-energy materials [98], NMR spin–spin coupling constants from the electron delocalization indices [99, 100], simultaneous consistent prediction of five bulk properties of liquid HF in MD simulation [101], classification of atom types in proteins with future potential applications in force-field design [60, 102–104], reconstructing large molecules from transferable fragments or atoms in molecules [60, 105–119] (see also Chapters 11 and 12), atomic partitioning of the molecular electrostatic potential [120–122], prediction of hydrogen-bond donor capacity [123] and basicity [124], and to provide an atomic basis for curvature-induced polarization in carbon nanotubes and nanoshells [125].

1.10 Steps of a Typical QTAIM Calculation

It should be clear from the outset that QTAIM applies equally well to experimental [2, 126] and calculated electron densities [1, 3]; in this tutorial, however, we will discuss calculated densities.

The starting point for the application of the QTAIM theory is the electron density. The density can be calculated from the many-electron single-determinant or many-determinant wavefunction (or Slater-like determinants built from Kohn–Sham orbitals in density functional theory [127]) obtained by a variety of methods and software. The electron density necessary for meaningful analysis by means of QTAIM must be obtained with a basis set flexible enough for an accurate representation of the bonding regions, in other words it must include polarization functions. In the case of anions, excited states, and weak bonding interactions between atoms separated by relatively large distances one must augment the basis set with diffuse functions. When heavy atoms are present in the molecule, which usually necessitates the use of effective core potentials (ECP), it is necessary to treat the valence shell and at least one sub-valence shell explicitly to obtain meaningful results from the integrations. It is important to note that bond paths cannot be traced to the nuclei of atoms described by ECPs. Alternatively, often the geometry is optimized with a basis set including the ECP on the heavy atoms followed by a single point calculation at the optimized geometry using a full basis set on all atoms.

The first step in a molecular QTAIM calculation is, thus, the generation of a wavefunction (or wavefunction-like single determinant in a DFT calculation [127]) from an electronic structure calculation with software such as Gaussian [61] or GAMESS [128].

The electron density derived from the wavefunction is then subjected to a point-by-point topological analysis to locate the bond critical points and the bond paths, by use of software such as EXTREME [129–131]. The space of the molecule is then partitioned by the zero-flux surfaces and atomic integrations are performed to obtain the atomic contributions to the molecular properties using software such as PROAIM and its variants [129–131]. EXTREME and PROAIM are both part of the AIMPAC suite of programs developed in Professor Bader's laboratory (*McMaster University*) [129–131].

Several other software packages derived from AIMPAC are available for analysis of the electron density according to QTAIM. Among the widely used programs are MORPHY [132], developed by Dr Paul Popelier's group (*University of Manchester*), and AIM2000 [133–135], developed by Professor Friederich Biegler–König (*University of Bielfeld*), both of which apply to molecular calculations. The program TOPOND [136] (described in Chapter 7) was developed by Dr Carlo Gatti (*National Research Council of Italy*) for the analysis of periodic densities obtained from CRYSTAL [137].

After all atomic integrations have been performed to the desired accuracy (as measured by the value of the integrated Laplacian) one typically uses shell scripts

and/or simple UNIX/Linux commands such as "grep" to extract the relevant information from the electronic integration files. In this manner the summarized results can be further imported to a spreadsheet or a plotting program. Integration files – which contain the atomic overlap matrices – can be subsequently analyzed by software such as AIMDELOC [44] or LI-DICALC [45, 46] to obtain the localization and delocalization indices.

Further, the wavefunction files can be used as input to plotting routines. GRDVEC can be used to generate two-dimensional plots of the gradient vector field and/or the interatomic surfaces and bond paths projected on a plane selected by the user (right half of Fig. 1.3a). Contour diagrams of the density (such as those in the left half of Fig. 1.3a, and Fig. 1.5), the Laplacian, or energy densities can be generated by first calculating the corresponding grid by the use of GRIDV software followed by the generation of the graphics file from the grid

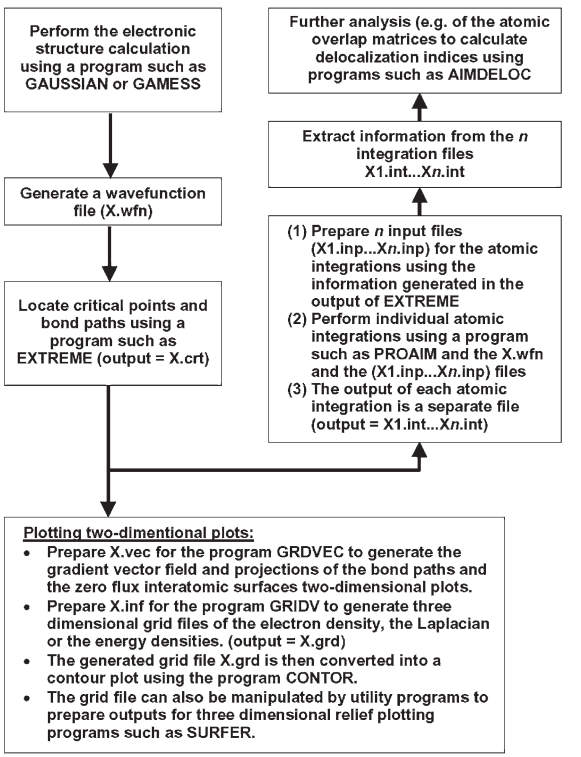

Perform the electronic structure calculation using a program such as GAUSSIAN or GAMESS
Generate a wavefunction file (X.wfn)
Locate critical points and bond paths using a program such as EXTREME (output = X.crt)
(1) Prepare n input files (X1.inp...Xn.inp) for the atomic integrations using the information generated in the output of EXTREME
(2) Perform individual atomic integrations using a program such as PROAIM and the X.wfn and the (X1.inp...Xn.inp) files
(3) The output of each atomic integration is a separate file (output = X1.int...Xn.int)
Extract information from the n integration files X1.int...Xn.int
Further analysis (e.g. of the atomic overlap matrices to calculate delocalization indices using programs such as AIMDELOC
Plotting two-dimentional plots:
• Prepare X.vec for the program GRDVEC to generate the gradient vector field and projections of the bond paths and the zero flux interatomic surfaces two-dimensional plots.
• Prepare X.inf for the program GRIDV to generate three dimensional grid files of the electron density, the Laplacian or the energy densities. (output = X.grd)
• The generated grid file X.grd is then converted into a contour plot using the program CONTOR.
• The grid file can also be manipulated by utility programs to prepare outputs for three dimensional relief plotting programs such as SURFER.

Fig. 1.6 The main steps in a simple QTAIM calculation. The software cited in the figure is part of the AIMPAC suite of programs [129–131]. Other programs are available that can perform most or all of these steps, including, for example, AIM2000 [133–135], MORPHY [132], and AIMALL97 [139].

by using a program such as CONTOR. (GRDVEC, GRIDV, and CONTOR are components of AIMPAC [129–131]).

The grid generated by GRIDV can be manipulated by utility programs such as GridV_REFORMATTER (available from the authors) to generate inputs for programs such as Surfer [138] which produce three-dimensional relief maps of the field represented by the calculated grid (Fig. 1.1b is an example).

The main steps of a typical QTAIM calculation are summarized in Fig. 1.6.

Appendix: The Inexact Satisfaction of the Molecular Virial Theorem in Electronic Structure Calculations

For a molecule in an equilibrium geometry (with vanishing forces on the nuclei), the molecular virial theorem is expressed as:

$$\gamma = -\frac{V}{T} = 2 \tag{45}$$

Because of the propagation of numerical errors, small (but non-vanishing) thresholds of convergence of both the SCF and the geometry optimization steps, and the use of incomplete basis sets, electronic structure calculations do not usually satisfy the virial theorem exactly and the virial ratio (γ) can deviate by perhaps as much as 0.01 from the ideal value of 2. As a result of this deviation, atomic energies will not sum to yield the molecular energy with acceptable accuracy.

Atomic integration software such as PROAIM [129–131] correct for this error numerically. Thus, instead of simply multiplying each atomic kinetic energy $T(\Omega)$ by (-1) to obtain the total atomic energy $E(\Omega)$, the latter is obtained by multiplying $T(\Omega)$ by $(1-\gamma)$. These corrected atomic energies do satisfy Eq. (37), and their sum equals the total molecular energy to within a small numerical *integration* error. The virial corrections usually scale linearly with regard to $T(\Omega)$ which, fortunately, leaves the relative stabilities of the atoms unchanged.

The integration software obtains the virial ratio from the wavefunction files generated by Gaussian [61] or GAMESS [128]. The virial is printed in the last line in the wavefunction file. For Hartree–Fock or density functional calculations, Gaussian prints the correct virial in the wavefunction file and the integrations proceed without problems. For wavefunction files calculated at a post Hartree–Fock level, for example those obtained using Møller–Plesset perturbation theory (MP*n*) or configuration interaction methods (CI), the virial printed in the wavefunction file generated using the Gaussian 98 or 03 [61] programs (which are available at the time of writing) is *the Hartree–Fock virial and not that of the current post-Hartree–Fock method* [even if the key word "DENSITY = CURRENT" is invoked and despite the fact that the correct (current) wavefunction is printed]. If such a wavefunction file is fed directly to an integration program, the calculated atomic energies will be rectified using the Hartree–Fock γ (instead of the post Hartree–Fock γ), resulting in atomic energies which do not add up to the molecular value.

In these circumstances the user must calculate the virial of the current method "by hand" from information contained in the Gaussian "log" or "out" output file [140], by dividing, for example, the MP2 (or other correlated total energy) by the kinetic energy listed just after the final electrical multipoles in the Gaussian output. The wavefunction files must then be edited to reflect this new "correct" virial before submitting it to the integration software [140].

In highly accurate calculations it is sometimes necessary to perform atomic integrations of energy densities obtained from systems which satisfy the molecular virial theorem exactly [16, 141]. The author of Chapter 3 of this book, Dr Todd A. Keith, has written a link [142] for Gaussian [61] implementing Löwdin's self-consistent virial scaling (SCVS) [143, 144] which produces final wavefunctions satisfying the virial theorem to a very high accuracy.

Acknowledgments

We thank Professor George Heard (University of North Carolina at Asheville) and Dr Katherine N. Robertson (Dalhousie University) for their helpful comments on this chapter, and Dr Todd Keith (Semichem, Inc.) and Dr Jamie Platts (Cardiff University) for discussions on the virial correction. The American Chemical Society is acknowledged for its permission to reproduce Fig. 1.5.

References

1 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press: Oxford, U.K., 1990.
2 P. Coppens, *X-ray Charge Densities and Chemical Bonding*, Oxford University Press, Inc.: New York, 1997.
3 P. L. A. Popelier, *Atoms in Molecules: An Introduction*, Prentice Hall: London, 2000.
4 C. F. Matta, N. Castillo, R. J. Boyd, *J. Phys. Chem. A* **2005**, *109*, 3669–3681.
5 N. Castillo, C. F. Matta, R. J. Boyd, *Chem. Phys. Lett.* **2005**, *409*, 265–269.
6 C. Gatti, P. Fantucci, G. Pacchioni, *Theor. Chem. Acc. (Formerly, Theoret. Chim. Acta)* **1987**, *72*, 433–458.
7 W. L. Cao, C. Gatti, P. J. MacDougall, R. F. W. Bader, *Chem. Phys. Lett.* **1987**, *141*, 380–385.
8 M. Sakata, *Acta Cryst. A* **1990**, *46*, 263–270.
9 R. Y. de Vries, W. J. Briels, D. Feil, G. te Velde, E. J. Baerends, *Can. J. Chem.* **1996**, *74*, 1054–1058.
10 A. Taylor, C. F. Matta, R. J. Boyd, submitted for publication **2006**.
11 R. F. W. Bader, J. A. Platts, *J. Chem. Phys.* **1997**, *107*, 8545–8553.
12 R. F. W. Bader, *Phys. Rev. B* **1994**, *49*, 13348–13356.
13 J. Schwinger, *Phys. Rev.* **1951**, *82*, 914–927.
14 R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102*, 7314–7323.
15 T. A. Keith, R. F. W. Bader, Y. Aray, *Int. J. Quantum Chem.* **1996**, *57*, 183–198.
16 C. F. Matta, J. Hernández-Trujillo, T. H. Tang, R. F. W. Bader, *Chem. Eur. J.* **2003**, *9*, 1940–1951.
17 C. F. Matta, Chapter 9 in: *Hydrogen Bonding – New Insight, (S. J. Grabowski, Ed.)*, Springer: 2006, pp 337–376.
18 R. F. W. Bader, C. F. Matta, *J. Phys. Chem. A* **2004**, *108*, 8385–8394.
19 R. P. Sagar, A. C. T. Ku, V. H. Jr. Smith, A. M. Simas, *J. Chem. Phys.* **1988**, *88*, 4367–4374.

20 Z. Shi, R. J. Boyd, *J. Chem. Phys.* **1988**, *88*, 4375–4377.

21 R. W. F. Bader, G. L. Heard, *J. Chem. Phys.* **1999**, *111*, 8789–8797.

22 R. J. Gillespie, R. S. Nyholm, *Quart. Rev. Chem. Soc.* **1957**, *11*, 339.

23 R. J. Gillespie, R. S. Nyholm, in: *Progress in Stereochemistry* (W. Klyne, P. B. D. de la Mare, Eds.), Butterworths: London, 1958.

24 R. J. Gillespie, I. Hargittai, *The VSEPR Model of Molecular Geometry*, Allyn and Bacon: Boston, 1991.

25 R. F. W. Bader, P. J. MacDougall, C. D. H. Lau, *J. Am. Chem. Soc.* **1984**, *106*, 1594–1605.

26 R. F. W. Bader, R. J. Gillespie, P. J. MacDougall, *J. Am. Chem. Soc.* **1988**, *110*, 7329–7336.

27 R. J. Gillespie, I. Bytheway, T.-H. Tang, R. F. W. Bader, *Inorg. Chem.* **1996**, *35*, 3954–3963.

28 M. T. Carroll, C. Chang, R. F. W. Bader, *Mol. Phys.* **1988**, *63*, 387–405.

29 M. T. Carroll, J. R. Cheeseman, R. Osman, H. Weinstein, *J. Phys. Chem.* **1989**, *93*, 5120–5123.

30 R. J. Boyd, S. C. Choi, *Chem. Phys. Lett.* **1986**, *129*, 62–65.

31 M. T. Carroll, R. F. W. Bader, *Mol. Phys.* **1988**, *65*, 695–722.

32 E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170–173.

33 S. J. Grabowski, *J. Phys. Chem. A* **2001**, *105*, 10739–10746.

34 M. Domagala, S. Grabowski, K. Urbaniak, G. Mloston, *J. Phys. Chem. A* **2003**, *107*, 2730–2736.

35 S. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2005**, *109*, 4331–4341.

36 M. Domagala, S. Grabowski, *J. Phys. Chem. A* **2005**, *109*, 5683–5688.

37 O. Knop, R. J. Boyd, S. C. Choi, *J. Am. Chem. Soc.* **1988**, *110*, 7299–7301.

38 J. L. Jules, J. R. Lombardi, *J. Mol. Struct. (Theochem)* **664–665**, *2003*, 255–271.

39 S. T. Howard, O. Lamarche, *J. Phys. Org. Chem.* **2003**, *16*, 133–141.

40 R. F. W. Bader, T. T. Nguyen-Dang, *Adv. Quantum Chem.* **1981**, *14*, 63–124.

41 D. Cremer, E. Kraka, *Angew. Chem. Int. Ed. Engl.* **1984**, *23*, 627–628.

42 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A* **1999**, *103*, 304–314.

43 R. F. W. Bader, M. E. Stephens, *J. Am. Chem. Soc.* **1975**, *97*, 7391–7399.

44 C. F. Matta, *AIMDELOC (QCPE 0802)* Quantum Chemistry Program Exchange, Indiana University, 2001. (http://qcpe.chem.indiana.edu/).

45 Y.-G. Wang, C. F. Matta, N. H. Werstiuk, *J. Comput. Chem.* **2003**, *24*, 1720–1729.

46 Y.-G. Wang, N. H. Werstiuk, *J. Comput. Chem.* **2003**, *24*, 379–385.

47 M. A. Austen, *A New Procedure for Determining Bond Orders in Polar Molecules, with Applications to Phosphorus and Nitrogen Containing Systems, Ph.D. Thesis*, McMaster University: Hamilton, Canada, 2003.

48 C. F. Matta, J. Hernández-Trujillo, *J. Phys. Chem. A* **2003**, *107*, 7496–7504 (Correction: *J. Phys. Chem. A* **2005**, *109*, 10798).

49 E. A. Zhurova, C. F. Matta, N. Wu, V. V. Zhurov, A. A. Pinkerton, *J. Am. Chem. Soc.* **2006**, *128*, 8849–8861.

50 R. F. W. Bader, P. F. Zou, *Chem. Phys. Lett.* **1992**, *191*, 54–58.

51 C. F. Matta, R. F. W. Bader, *J. Phys. Chem. A* **2006**, *110*, 6365–6371.

52 (a) L. Cohen, *J. Chem. Phys.* **1979**, *70*, 788–789; (b) L. Cohen, *J. Chem. Phys.* **1984**, *80*, 4277–4279.

53 F. Cortés-Guzmán, J. Hernández-Trujillo, G. Cuevas, *J. Phys. Chem. A* **2003**, *107*, 9253–9256.

54 R. F. W. Bader, J. R. Cheeseman, K. E. Laidig, K. B. Wiberg, C. Breneman, *J. Am. Chem. Soc.* **1990**, *112*, 6530–6536.

55 R. F. W. Bader, C. F. Matta, *Int. J. Quantum Chem.* **2001**, *85*, 592–607.

56 C. F. Matta, *Applications of the Quantum Theory of Atoms in Molecules to Chemical and Biochemical Problems, Ph.D. Thesis*, McMaster University: Hamilton, Canada, 2002. (Available on line, http://chem.utoronto.ca/~cmatta/).

57 C. F. Matta, R. J. Gillespie, *J. Chem. Educ.* **2002**, *79*, 1141–1152.

58 L. Pauling, *The Nature of the Chemical Bond, (Third Ed.)*, Cornell University Press: Ithaca, N.Y., 1960.
59 C. F. Matta, *FRAGDIP (QCPE 0801)* Quantum Chemistry Program Exchange, Indiana University, 2001. (http://qcpe.chem.indiana.edu/).
60 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2003**, *52*, 360–399.
61 M. J. Frisch, *et al.*, Gaussian 03, Gaussian Inc.: Pittsburgh PA, 2003.
62 C. G. Gray, K. E. Gubbins, *Theory of Molecular Fluids, (Vol. 1)*, Clarendon Press: Oxford, 1984.
63 E. Espinosa, E. Molins, *J. Chem. Phys.* **2000**, *113*, 5686–5694.
64 O. A. Zhikol, O. Shishkin, K. A. Lyssenko, J. Leszczynski, *J. Chem. Phys.* **2005**, *122*, 144104-1–144104-8.
65 M. P. Waller , A. Robertazzi, J. A. Platts, D. E. Hibbs, P. A. Williams, *J. Comput. Chem.* **2006**, *27*, 491–504.
66 C. F. Matta, N. Castillo, R. J. Boyd, *J. Phys. Chem. B* **2006**, *110*, 563–578.
67 S. E. O'Brien, P. L. A. Popelier, *Can. J. Chem.* **1999**, *77*, 28–36.
68 P. L. A. Popelier, *J. Phys. Chem. A* **1999**, *103*, 2883–2890.
69 S. E. O'Brien, P. L. A. Popelier, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 764–775.
70 U. A. Chaudry, P. L. A. Popelier, *J. Org. Chem.* **2004**, *69*, 233–241.
71 R. Carbó, L. Leyda, M. Arnau, *Int. J. Quantum Chem.* **1980**, *17*, 1185–1189.
72 K. B. Wiberg, R. F. W. Bader, C. D. H. Lau, *J. Am. Chem. Soc.* **1987**, *109*, 1001–1012.
73 T. A. Keith, R. F. W. Bader, *Chem. Phys. Lett.* **1992**, *194*, 1–8.
74 T. A. Keith, R. F. W. Bader, *Int. J. Quantum Chem.* **1996**, *60*, 373–379.
75 R. F. W. Bader, T. A. Keith, *J. Chem. Phys.* **1993**, *99*, 3683–3693.
76 T. A. Keith, R. F. W. Bader, *Chem. Phys. Lett.* **1993**, *210*, 223–231.
77 T. A. Keith, R. F. W. Bader, *J. Chem. Phys.* **1993**, *99*, 3669–3682.
78 R. F. W. Bader, M. T. Carroll, J. R. Cheeseman, C. Chang, *J. Am. Chem. Soc.* **1987**, *109*, 7968–7979.
79 R. F. W. Bader, T. A. Keith, K. M. Gough, K. E. Laidig, *Mol. Phys.* **1992**, *75*, 1167–1189.
80 K. E. Laidig, *Can. J. Chem.* **1996**, *74*, 1131–1138.
81 K. M. Gough, M. M. Yacowar, R. H. Cleve, J. R. Dwyer, *Can. J. Chem.* **1996**, *74*, 1139–1144.
82 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Chem. Phys.* **1993**, *98*, 9669–9677.
83 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Phys. Chem.* **1994**, *98*, 771–776.
84 K. M. Gough, H. K. Srivastava, *J. Phys. Chem.* **1996**, *100*, 5210–5216.
85 R. L. A. Haiduke, A. E. de Oliveira, R. E. Bruns, *J. Phys. Chem. A* **2004**, *108*, 6788–6796.
86 J. V. da Silva, R. L. A. Haiduke, R. E. Bruns, *J. Phys. Chem. A* **2006**, *110*, 4839–4845.
87 P. H. César, S. H. D. M. Faria, J. V. da Silva Jr., R. L. A. Haiduke, R. E. Bruns, *Chem. Phys.* **2005**, *317*, 35–42 .
88 R. L. A. Haiduke, R. E. Bruns, *J. Phys. Chem. A* **2005**, *109*, 2680–2688.
89 R. F. W. Bader, D. Bayles, G. L. Heard, *J. Chem. Phys.* **2000**, *112*, 10095–10105.
90 K. E. Laidig, *Chem. Phys. Lett.* **1991**, *185*, 483–489.
91 R. F. W. Bader, *Mol. Phys.* **2002**, *100*, 3333–3344.
92 P. F. Zou, R. F. W. Bader, *Acta Cryst. A* **1994**, *50*, 714–725.
93 R. F. W. Bader, D. Bayles, *J. Phys. Chem. A* **2000**, *104*, 5579–5589.
94 R. F. W. Bader, A. Streitwieser, A. Neuhaus, K. E. Laidig, P. Speers, *J. Am. Chem. Soc.* **1996**, *118*, 4959–4965.
95 R. F. W. Bader, S. Johnson, T.-H. Tang, P. L. A. Popelier, *J. Phys. Chem.* **1996**, *100*, 15398–15415.
96 K. R. Adam, *J. Phys. Chem. A* **2002**, *106*, 11963–11972.
97 M. Song, C. M. Breneman, J. Bi, N. Sukumar, K. P. Bennett, S. Cramer, N. Tugcu, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1347–1357.
98 C. M. Breneman, M. Rhem, *J. Comput. Chem.* **1997**, *18*, 182–197.
99 C. F. Matta, J. Hernández-Trujillo, R. F. W. Bader, *J. Phys. Chem. A* **2002**, *106*, 7369–7375.

100 N. Castillo, C. F. Matta, R. J. Boyd, *J. Chem. Inf. Mod.* **2005**, *45*, 354–359.

101 S. Y. Liem, P. L. A. Popelier, *J. Chem. Phys.* **2003**, *119*, 4560–4566.

102 H. J. Bohórquez, M. Obregón, C. Cárdenas, E. Llanos, C. Suárez, J. L. Villaveces, M. E. Patarroyo, *J. Phys. Chem. A.* **2003**, *107*, 10090–10097.

103 P. L. A. Popelier, F. M. Aicken, *ChemPysChem* **2003**, *4*, 824–829.

104 P. L. A. Popelier, F. M. Aicken, *J. Am. Chem. Soc.* **2003**, *125*, 1284–1292.

105 C. Chang, R. F. W. Bader, *J. Phys. Chem.* **1992**, *96*, 1654–1662.

106 C. M. Breneman, T. R. Thompson, M. Rhem, M. Dung, *Comput. Chem.* **1995**, *19*, 161–179.

107 C. M. Breneman, L. W. Weber, in: *The application of charge density research to chemistry and drug design (NATO ASI Series)*, (G. A. Jeffrey, J. F. E. Piniella, Eds.) Plenum Press, New York, 1991, pp 357–358.

108 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2000**, *40*, 310–329.

109 C. F. Matta, R. F. W. Bader, *Proteins: Struct. Funct. Genet.* **2002**, *48*, 519–538.

110 R. F. W. Bader, F. J. Martín, *Can. J. Chem.* **1998**, *76*, 284–291.

111 F. J. Martín, *Theoretical Synthesis of Macromolecules from Transferable Functional Groups, Ph.D. Thesis*, McMaster University: Hamilton, 2001.

112 R. F. W. Bader, C. F. Matta, F. J. Martín, Chapter 7 in: *Medicinal Quantum Chemistry*, (F. Alber, P. Carloni, Eds.) Wiley–VCH: Weinheim, 2003, pp 201–231.

113 V. Pichon-Pesme, C. Lecomte, R. Wiest, M. Bénard, *J. Am. Chem. Soc.* **1992**, *114*, 2713–2715.

114 R. Wiest, V. Pichon-Pesme, M. Bénard, C. Lecomte *J. Phys. Chem.* **1994**, *98*, 1351–1362.

115 V. Pichon-Pesme, C. Lecomte, H. Lachekar, *J. Phys. Chem.* **1995**, *99*, 6242–6250.

116 C. Jelsch, V. Pichon-Pesme, C. Lecomte, A. Aubry, *Acta Cryst. D* **1998**, *54*, 1306–1318.

117 C. F. Matta, *J. Phys. Chem. A* **2001**, *105*, 11088–11101.

118 B. Dittrich, T. Koritsánszky, M. Grosche, W. Scherer, R. Flaig, A. Wagner, H. G. Krane, H. Kessler, C. Riemer, A. M. M. Schreurs, P. Luger, *Acta Cryst. B* **2002**, *58*, 721–727.

119 S. Scheins, M. Messerschmidt, P. Luger, *Acta Cryst. B* **2005**, *61*, 443–448.

120 D. S. Kosov, P. L. A. Popelier, *J. Chem. Phys.* **2000**, *113*, 3969–3974.

121 D. S. Kosov, P. L. A. Popelier, *J. Phys. Chem. A* **2000**, *104*, 7339–7345.

122 P. L. A. Popelier, L. Joubert, D. S. Kosov, *J. Phys. Chem. A* **2001**, *105*, 8254–8261.

123 J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 973–980.

124 J. A. Platts, *Phys. Chem. Chem. Phys.* **2000**, *2*, 3115–3120.

125 T. Dumitrica, C. M. Landis, B. I. Yakobson, *Chem. Phys. Lett.* **2002**, *360*, 182–188.

126 T. S. Koritsanszky, P. Coppens, *Chem. Rev.* **2001**, *101*, 1583–1628.

127 W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory, (Second Edition)*, Wiley–VCH: New York, 2001.

128 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.* **1993**, *14*, 1347–1363.

129 Bader, R. F. W., http://www.chemistry.mcmaster.ca/aimpac/.

130 F. W. Biegler-König, T. T. Nguyen-Dang, Y. Tal, R. F. W. Bader, A. J. Duke *J. Phys. B: At. Mol. Phys.* **1981**, *14*, 2739–2751.

131 F. W. Biegler-König, R. F. W. Bader, T.-H. Tang, *J. Comput. Chem.* **1982**, *13*, 317–328.

132 Popelier, P. L. A., MORPHY, UMIST, England, EU, **1998**.

133 F. W. Biegler-König, J. Schönbohm, D. Bayles, *J. Comput. Chem.* **2001**, *22*, 545–559.

134 F. W. Biegler-König, *J. Comput. Chem.* **2000**, *21*, 1040–1048.

135 F. W. Biegler-König, J. Schönbohm, D. Bayles, AIM2000 Website: http://gauss.fh-bielefeld.de/aim2000.

136 C. Gatti, TOPOND, CNR-CSRSRC: Milano, 1998.

137 V. R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, N. M. Harrison, K. Doll, B. Civalleri, I. J. Bush, Ph. D'Arco, M. Llunell, *CRYSTAL* 2003.

138 Surfer 08, Golden Software, Inc., Golden, Colorado, USA, 2002.

139 T. A. Keith, *AIMALL97 for Windows*, 1997.

140 J. A. Platts, CCL: MP2 virial value (8 July 2005), http://www.ccl.net/cgi-bin/ccl/message.cgi?2005+07+08+004.

141 F. Cortés-Guzmán, R. F. W. Bader, *Chem. Phys. Lett.* **2003**, *379*, 183–192.

142 T. A. Keith, Link for self-consistent virial scaling (SCVS) in Gaussian 94/98, **1998**.

143 P.-O. Löwdin, *J. Mol. Spectr.* **1959**, *3*, 46–66.

144 D. E. Magnoli, J. R. Murdoch, *Int. J. Quantum Chem.* **1982**, *22*, 1249–1262.

Part I
Advances in Theory

2
The Lagrangian Approach to Chemistry

Richard F. W. Bader

2.1
Introduction

This article is written by a chemist for other chemists to explain how and why the definition of an atom in a molecule requires that one foregoes the usual Hamiltonian approach to quantum mechanics and replace it with one expressed in terms of the Lagrangian and the associated action principle. A consequence of the Lagrangian approach to chemistry is the possibility of foregoing arbitrary models and dealing instead with observation and physics, so the article should be of particular interest to experimentalists. It is addressed in particular to younger chemists, those willing to extend their knowledge of quantum mechanics beyond the orbital model of electronic structure.

2.1.1
From Observation, to Physics, to QTAIM

Thirty-four years have passed since it was first postulated that the virial theorem should apply to a bounded region of real space – to an atom in a molecule [1]. This postulate was put forth on the basis of the observation of the paralleling transferability of the topological properties of the electron and kinetic energy densities. In another three years, bringing one to 1975, this postulate was derived starting from Schrödinger's first paper on "wave mechanics" [2]. The 1975 paper, by yielding a variational derivation of the virial theorem for an atom in a molecule, was the initial step in establishing a *variational definition* of "an atom in a molecule", one derived from the Lagrangian approach to physics. Thus one proceeded from observation, to physics, to the quantum theory of atoms in molecules, QTAIM, a theory that applies equally to the total system and to its constituent atoms [3]. By building upon Schwinger's principle of stationary action [4] that follows the Lagrangian approach to physics, QTAIM yields a variational statement of the Heisenberg equation of motion for all quantum observables and establishes the equal applicability of these equations to the total system and its con-

stituent atoms. In the Heisenberg approach, quantum mechanics is represented in terms of the equations of motion of the observables, thus emphasizing the ties that link theory to observation. Thus QTAIM, by uniting the Heisenberg and Lagrangian approaches to physics, unites theoretical chemistry with experiment.

QTAIM brings to the fore the modus operandi of quantum mechanics – to use the equations of motion – the theorems of quantum mechanics – to predict and understand the observed properties of a system. This is the approach followed by Slater in his use of the virial and Feynman theorems which he considered to be "two of the most powerful theorems applicable to molecules and solids" [5]. QTAIM, by extending these and all theorems to an atom in a molecule, enables one to apply this approach to all problems at the atomic level. The time has arrived to replace the criticisms of these powerful theorems that one finds throughout the literature, with the clarity and predictive powers of physics.

2.2
The Lagrangian Approach

2.2.1
What is The Lagrangian Approach and What Does it Do?

The Hamiltonian approach to quantum mechanics evolved from Schrödinger's wave equation that he derived in 1926 [6]. Yourgrau and Mandelstam (Y&M) in their prelude to introducing the Lagrangian approach in their monograph on variational principles in quantum theory state: "The ordinary Hamiltonian presentation of quantum mechanics contains, apart from the basic principles of the subject, two distinct postulates: the commutation relations between generalized co-ordinates and the momenta, and the equations of motion" [7]. The alternative "Lagrangian approach", is based on the action principle, which in its earliest form was referred to as *the principle of least action* – first enunciated somewhat imperfectly by Maupertuis in 1744 and stated in its present form by Hamilton in 1834. Y&M point out that in classical mechanics the Lagrangian formalism can be substituted for the Hamilton theory enabling the laws of mechanics to be expressed as *a single postulate*, the principle of least action stated by Hamilton. Y&M go on to ask: "The obvious question that arises from such considerations: is it not possible, and moreover most desirable, to obtain a similar Lagrangian formulation of the quantum laws depending upon a single postulate only? This aim is achieved by the Feynman and Schwinger theories" [7]. Their theories re-introduced the action principle into the active fabric of physics, Feynman [8] in 1948 and Schwinger [4] in 1951, and yield both the equations of motion and the commutation relations. This chapter shows how Schwinger's formulation enables one to ask and answer the question "what is an atom in a molecule?", a question that cannot be posed within the Hamiltonian approach. Answering the question takes one on an intellectual journey beginning with Schrödinger, on to Dirac, and end-

ing with Schwinger, whose principle of stationary action can be extended to yield the physics of an open system and thus, the theory of atoms in molecules [3]. The work of Feynman and Schwinger may provide "the real foundation of quantum mechanics" [9]. Its extension to an atom in a molecule will bring to bear the advantages of this approach to the understanding and prediction of chemistry [10].

The derivation of the atomistic approach from the action principle imparts the certainty of physics to the molecular structure hypothesis – that a molecule is a collection of atoms with additive, characteristic properties linked by a network of bonds that impart a structure – a concept forged in the crucible of nineteenth-century experimental chemistry. One hundred and fifty years of *experimental* chemistry underlie the realization that the properties of a total system are the sum of its atomic contributions and Dalton's atomic hypothesis has emerged as the operational theory of chemistry. The time has arrived for a shift of emphasis in the prediction and understanding of chemistry, away from a philosophy based on the premise held by some that because chemical concepts are beyond physics, one is free to formulate and choose between a plethora of frequently conflicting subjective models.

This chapter is intended for a wide audience. Each section opens with an introduction to the essential underlying ideas and their reading will take one through to the chapter's goal – an understanding of the rooting of the atomic concept in the fabric of quantum mechanics. Experimental chemists with a limited knowledge of physics, such as I when I first switched to theory, may skip intervening sections that provide those more knowledgeable of physics with more mathematical detail. A recent article provides a narrative of the development of the theory of atoms in molecules from the properties of the measurable electron density distribution [11].

2.2.2
The Lagrangian and the Action Principle – A Return to the Beginnings

It was the classical concept of "action" that provided Schrödinger with the conceptual basis for his derivation of $\hat{H}\psi = E\psi$ in the first of the four papers on "wave mechanics" that he published in 1926 [6]. He based its derivation on an analogy with a method of obtaining a solution to the classical equations of motion that corresponds to finding the time integral of the Lagrangian over the motion in question, the "action integral". Thus the action provided the starting point for Schrödinger's derivation of his equation that evolved into the Hamiltonian approach to quantum mechanics. The quantum theory of an atom in a molecule (QTAIM) was, in turn, first obtained by extension of Schrödinger's derivation to a system with finite boundaries [2]. Thus the initial step on the path that leads to the development an atom in a molecule is accomplished by a return to the action at the beginning of quantum mechanics, a path that leads directly to Schwinger's principle of stationary action [4]. It is the purpose of this section to demonstrate that wave mechanics had its origin in the action principle and to show

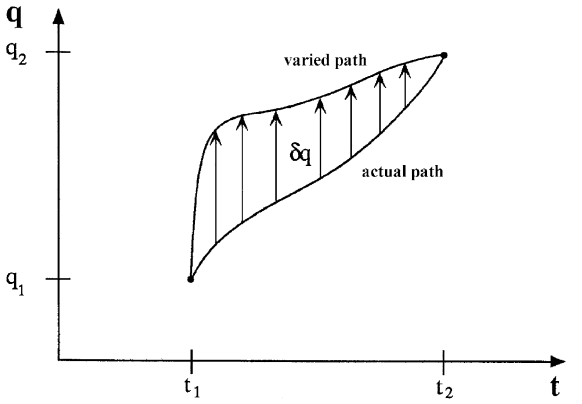


Fig. 2.1 A representation of the actual path and a possible varied path obtained by variation of the position coordinate q connecting the initial (q_1, t_1) and final (q_2, t_2) space–time points for a classical system.

how this leads unavoidably to the definition of an atom as a bounded piece of real space.

2.2.3
Minimization of the Action

The principle of least action states that a quantity called the action is minimized as a system moves from one configuration to another – the action is said to be "stationary" relative to variations in the space–time path connecting the two configurations. Such a path is depicted in Fig. 2.1 for a classical trajectory connecting initial and final points in space; q_1 and q_2, at the corresponding times t_1 and t_2.

The action, denoted by $\mathcal{W}_{12}$, with the dimensions of Planck's constant h (energy × time) is the time integral between the limits t_1 and t_2, of the Lagrangian $\mathcal{L}(q, \dot{q}, t)$, a function of the coordinates q, their velocities $\dot{q}$ and the time t (Eq. 1):

$$\mathcal{W}_{12} = \int_{t1}^{t2} \mathcal{L}(q, \dot{q}, t)\, dt \tag{1}$$

The mathematical procedure used to minimize the action over an entire path is outlined below for those desirous of a more complete description. The procedure yields a differential equation, the Euler equation, the equation of motion for the chosen Lagrangian; the classical action integral yielding Lagrange's equations and the quantum action yielding Schrödinger's equation. Thus the action principle enables one to *derive* Newton's and Schrödinger's equations and in this sense alone it is a more fundamental statement of physics.

2.2.4 Steps in Minimizing the Action

The classical Lagrangian equals the difference between the kinetic and potential energies, $\mathscr{L} = T - V$. It has a value at each point on the trajectory and the "sum" of these values between the two time limits must be a minimum. This problem is distinct from finding an extremum in some function at a single point in space using differential calculus. Determining an extremum of a function over an entire path is accomplished using the "calculus of variations". Assuming the existence of the "actual path" that is to be found, one generates a "varied path" coterminus in space and time, by displacing q on the actual path by an amount δq at each time t, as depicted in Fig. 2.1, thereby causing a variation or first-order change in the action integral, $\delta \mathscr{W}_{12}$. Because $\mathscr{W}_{12}$ must be a minimum for the true path, the variation $\delta \mathscr{W}_{12}$ must vanish, and the action is said to be "stationary". One of Feynman's lectures [12] presents a clear and very readable presentation of the elementary mathematics underlying the derivation of the expression for $\delta \mathscr{W}_{12} = 0$.

The mathematical result of varying the action integral is shown in Eq. (2) for a single coordinate q.

$$\delta \mathscr{W}_{12} = \int_{t1}^{t2} \{(\partial \mathscr{L}/\partial q) - d(\partial \mathscr{L}/\partial \dot{q})/dt\} \delta q \, dt = 0 \tag{2}$$

The variations in $\dot{q}$ are re-expressed in terms of δq using an integration by parts (refer to Feynman's lecture) and the resulting terms at the time end points are discarded. Thus, the result of the variation is given by the group of terms enclosed in the curly brackets in Eq. (2), all multiplied only by δq, the variations in q. Because δq is arbitrary, the only way in which the variation $\delta \mathscr{W}_{12}$ can vanish is for the group of terms in the curly brackets to equal zero, yielding a differential equation, representing Lagrange's equations:

$$\partial \mathscr{L}/\partial q - d(\partial \mathscr{L}/\partial \dot{q})/dt = 0 \tag{3}$$

This, as pointed out above, is a general result – minimizing the action generates a differential equation, called the Euler equation that for the classical action is Lagrange's equation of motion. For a single particle $\mathscr{L} = m\dot{q}^2/2 - V(q)$ and the equation reduces to Newton's equation of motion, i.e. the force, given by $-\partial V/\partial q$, equals mass times acceleration, $m\ddot{q}$.

If one employs the quantum mechanical Lagrangian and causes the action to be stationary relative to first-order variations in the wave function Ψ, one obtains Schrödinger's time dependent equation, $i\hbar \partial \Psi/\partial t = \hat{H}\Psi$, as the Euler equation. Because $\Psi(q, t)$ is a function of the coordinates and time, one must perform the variations over the whole of configuration space – all values of q – between the two time limits, a procedure that again clearly requires the methods of the calculus of variations.

2.3 The Action Principle in Quantum Mechanics

2.3.1 Schrödinger's Appeal to the Action

The purpose of this section is to show how Schrödinger was led to his derivation of $\hat{H}\psi = E\psi$ by a procedure that is equivalent to the variation of a constrained action integral for an infinitesimal time interval. Schrödinger derived $\hat{H}\psi = E\psi$ by constructing an expression for the energy of a quantum system in a stationary state expressed in terms of a "wave function $\psi(q)$" whose constrained variation over the whole of configuration space using the calculus of variations yields $\hat{H}\psi = E\psi$ as the Euler equation. He based the form of his functional on the classical Hamilton–Jacobi (H–J) equation whose solution $S(q,t)$ is the indefinite time integral of the Lagrangian – the action integral.

For most of classical mechanics – to solve Hamilton's equations of motion, for example – the Hamiltonian is expressed in terms of the coordinates q and the momenta p, $H[q,p]$. The Hamiltonian in the H–J equation expresses the momentum in terms of $\partial S/\partial q$, however, and the H–J equation for a time-independent system is of the form $H[q,\partial S/\partial q] = E$. Schrödinger argued to replace the solution $S(q)$ by an unknown function $\psi(q)$ such that the H–J equation be expressed as $H[q,(\hbar/\psi)\partial\psi/\partial q] = E$. Schrödinger did not seek solutions to this equation as in the classical case, but instead chose to "seek a function ψ such that for any arbitrary variation of it the integral of the said quadratic form, taken over the whole coordinate space, is stationary, ..." Then in his italics "*The quantum conditions are replaced by this variation problem*" [6].

The result of Schrödinger basing his energy integral on the H–J equation is that the momentum is expressed as a spatial derivative of the wave function, a term proportional to $\partial\psi/\partial q$ and, as a consequence, Schrödinger's energy integral is, aside from a trivial change in sign and before the imposition of the normalization constraint, equal to the quantum Lagrangian for a stationary state. Thus Schrödinger's variation to obtain $\hat{H}\psi = E\psi$ as the Euler equation is equivalent to the variation of a constrained action integral for an infinitesimal time interval.

2.3.2 Schrödinger's Minimization

This section summarizes the steps in the minimization procedure that lead to the appearance of the Hamiltonian operator and to $\hat{H}\psi = E\psi$, the details of the derivation being given in my book [3]. The energy expression that Schrödinger minimized by varying ψ to obtain his equation is given in Eq. (4). Schrödinger did not allow ψ to be complex in his first paper. The integrand is, by analogy with the Hamiltonian in the H–J equation, a function of $\psi(q)$ and $\partial\psi/\partial q$.

$$\mathscr{G}(\psi,\nabla\psi)=\int_{-\infty}^{+\infty}\mathrm{d}\mathbf{r}\{(\hbar^2/2\mathrm{m})((\partial\psi/\partial\mathrm{x})^2 + (\partial\psi/\partial\mathrm{y})^2+(\partial\psi/\partial\mathrm{z})^2)+(\mathrm{V}-\mathrm{E})\psi\psi\} \tag{4}$$

The first group of terms represents the kinetic energy, equivalently expressed as $(\hbar^2/2\mathrm{m})\nabla\psi\bullet\nabla\psi$ and V is the potential energy, set equal to $-\mathrm{e}^2/\mathrm{r}$ for the hydrogen atom in paper I. E is introduced into the expression as an undetermined multiplier to ensure that the wave function remain normalized and is identified with the total energy in the Euler equation.

The variation of $\mathscr{G}(\psi,\nabla\psi)$ is actually quite straightforward. If $\mathrm{f}(\psi,\nabla\psi)$ denotes the integrand in Eq. (4), the variation is given by $\delta\mathscr{G}(\psi,\nabla\psi)=\int\mathrm{d}\mathbf{r}\{(\partial\mathrm{f}/\partial\psi)\delta\psi+(\partial\mathrm{f}/\partial\nabla\psi)\delta\nabla\psi\}$, where $\mathrm{f}(\psi,\nabla\psi)$ is varied with respect to both ψ and $\nabla\psi$. To obtain the Euler equation one must clear the expression of the term involving $\delta\nabla\psi$ a step accomplished using an integration by parts employing the identity $\nabla\psi\bullet\delta\nabla\psi=-\nabla^2\psi\delta\psi+\nabla\bullet(\nabla\psi\delta\psi)$. This step introduces the usual kinetic energy operator $-(\hbar^2/2\mathrm{m})\nabla^2\psi$ into the variational integral. The resulting expression for the variation is given in Eq. (5).

$$\delta\mathscr{G}(\psi,\nabla\psi)=\int\mathrm{d}\mathbf{r}\{\hat{\mathrm{H}}\psi-\mathrm{E}\psi\}\delta\psi + \oint\mathrm{dS}(\mathbf{r}_\mathbf{s})\{(\hbar^2/2\mathrm{m})\nabla\psi\bullet\mathbf{n}(\mathbf{r})\delta\psi\}=0 \tag{5}$$

and the integration by parts leads to the appearance of the Hamiltonian operator $\hat{\mathrm{H}}=(-\hbar^2/2\mathrm{m})\nabla^2+\mathrm{V}$. The variation consists of a contribution from the variation of ψ over the entire system all multiplied by $\delta\psi$ and another from its variation on the surface bounding the system. The surface term arises by applying Gauss' theorem to the final term of the identity used to rid the expression of $\nabla\delta\psi$. Because the surface in this case resides at infinity where, as Schrödinger points out, one requires $\delta\psi=0$, the surface term vanishes and the variation $\delta\mathscr{G}(\psi,\nabla\psi)$ reduces to:

$$\delta\mathscr{G}(\psi,\nabla\psi)=\int\mathrm{dr}\{\hat{\mathrm{H}}\psi-\mathrm{E}\psi\}\delta\psi=0 \tag{6}$$

The variation in Eq. (6) will vanish for arbitrary $\delta\psi$ only if $\hat{\mathrm{H}}\psi-\mathrm{E}\psi=0$. Thus, making $\mathscr{G}(\psi,\nabla\psi)$ stationary yields Schrödinger's equation for a stationary state [13].

2.3.2.1 Two Ways of Expressing the Kinetic Energy

$\mathscr{G}(\psi,\nabla\psi)$ expresses the kinetic energy in the form $+(\hbar^2/2\mathrm{m})\nabla\psi\bullet\nabla\psi$ rather than as $-(\hbar^2/2\mathrm{m})\psi\nabla^2\psi$, as it appears when obtained from Schrödinger's equation [14]. It is a consequence of the presence of this term that the extension of the variation of the action to an open system leads to its unique definition in terms

of the zero-flux boundary condition [2, 3]. The difference between the two forms of the kinetic energy is locally proportional to the Laplacian of the electron density ρ, as given without loss of generality in Eq. (7) for a one-electron system [3, 15]:

$$-(\hbar^2/4m)(\psi^*\nabla^2\psi + \psi\nabla^2\psi^*) - (\hbar^2/2m)\nabla\psi^* \bullet \nabla\psi = -(\hbar^2/4m)\nabla^2\rho \tag{7}$$

a result alternatively expressed as $K(\mathbf{r}) - G(\mathbf{r}) = L(\mathbf{r})$. Integration of Eq. (7) over a region of space Ω bounded by a surface $S(\Omega; \mathbf{r}_s)$, yields:

$$\begin{aligned} K(\Omega) - G(\Omega) &= -(\hbar^2/4m)\int_\Omega \nabla^2\rho(\mathbf{r})\,d\tau \\ &= -(\hbar^2/4m)\oint dS(\Omega;\mathbf{r})\nabla\rho(\mathbf{r}) \bullet \mathbf{n}(\mathbf{r}) \end{aligned} \tag{8}$$

identifying the average kinetic energies as $K(\Omega)$ and $G(\Omega)$, respectively. The volume integral of $\nabla^2\rho = \nabla \bullet \nabla\rho$, the divergence of a vector in Eq. (8), is replaced by the surface integral of the flux in $\nabla\rho$ through the surface of the region Ω using Gauss' theorem. When the region Ω refers to all space then $K(\Omega) = G(\Omega)$, because of the vanishing of ρ and its gradients at infinity, but for a region with finite boundaries the two quantities differ by the flux in $\nabla\rho$ through the surface of Ω and the kinetic energy is ill-defined. If, however, the surface $S(\Omega)$ is one of zero-flux in $\nabla\rho$ as defined in Eq. (9):

$$\nabla\rho(\mathbf{r}) \bullet \mathbf{n}(\mathbf{r}) = 0 \quad \text{for all points } \mathbf{r}_s \text{ on the surface } S(\Omega;\mathbf{r}_s) \tag{9}$$

where $\mathbf{n}(\mathbf{r})$ is a unit vector normal to the surface, then $K(\Omega) = G(\Omega)$ and the kinetic energy is a well defined quantity [1]. Equation (9) defines a surface that is not crossed by any trajectories traced out by the vector $\nabla\rho(\mathbf{r})$ and is consequently referred to as a "zero-flux surface". It is this surface that forms the boundary condition for an open quantum system when the action principle is extended to an open system [2, 3].

2.3.3
Obtaining an Atom from Schrödinger's Variation

What would happen if one were to repeat Schrödinger's derivation of his wave equation for an atom "in a molecule" rather than for an isolated atom? If one wishes to follow Schrödinger in the search for an atom in a molecule, one has no choice but to replace the infinite boundaries he placed on his energy expression in Eq. (4) by a set of finite limits that define a region Ω. That is, the atom will be defined as a bounded region of real space [2, 3]. The energy expression obtained by placing finite limits on $\mathcal{G}(\psi, \nabla\psi)$ in Eq. (4) to define a region Ω is denoted by $\mathcal{G}(\psi, \nabla\psi; \Omega)$. ($\psi$ is now allowed to be complex, with the variations in

ψ and ψ^* being independent). Its variation must include a term corresponding to a variation of the boundary if it is to be determined by the variational procedure rather than being arbitrarily assigned. The variation of $\mathscr{G}(\psi, \nabla\psi; \Omega)$ yields the same terms as given in Eq. (5) for $\delta\mathscr{G}(\psi, \nabla\psi)$ together with a second contribution to the surface integral from the variation of the surface. Because $\hat{\mathrm{H}}\psi = \mathrm{E}\psi$ still applies to the total system of which Ω is a part, the volume integral of the term $(\hat{\mathrm{H}}\psi - \mathrm{E}\psi)\delta\psi$ vanishes and only the surface terms remain in the variation of $\mathscr{G}(\psi, \nabla\psi; \Omega)$, Eq. (10):

$$\delta\mathscr{G}(\psi, \nabla\psi; \Omega) = \oint \mathrm{dS}(\Omega; \mathbf{r}_s)\{(\hbar^2/2\mathrm{m})\nabla\psi^* \bullet \mathbf{n}(\mathbf{r}_s)\delta\psi + \delta\mathrm{S}(\Omega; \mathbf{r}_s)\mathrm{f}(\psi, \nabla\psi)\} + \mathrm{cc} \tag{10}$$

The first surface term, as detailed in the variation of $\mathscr{G}(\psi, \nabla\psi)$, arises by ridding the expression of variations in $\nabla\psi$. The second term is the result of varying the surface of Ω by varying ψ. The function $\mathrm{f}(\psi, \nabla\psi)$ denotes the integrand in Schrödinger's functional, Eq. (4), and when evaluated in the surface and multiplied by the infinitesimal shift in the surface $\delta\mathrm{S}(\Omega)$, the term gives the contribution to the variation resulting from the variation of ψ on the surface. Because functions are not necessarily Hermitian over an open system Ω, the variation must include contributions from the complex conjugate (cc) terms. Equation (10), which does not define any particular surface, does not seem promising. Clearly the requirement of the principle of least action that the variation vanish must be discarded and the concept of stationarity broadened. This result is initially quite surprising if one lacks the knowledge that it forms the basis for the generalization of the principle of least action introduced by Schwinger.

2.3.3.1 The Role of Laplacian in the Definition of an Atom

If an atom is to be a bounded region of space, one might expect its physical description would require the presence of a surface integral that would describe contributions to its properties arising from the flux in currents through its surface. It is the purpose of this section to demonstrate that imposition of the zero-flux boundary condition, Eq. (9), during the variation of $\mathscr{G}(\psi, \nabla\psi; \Omega)$ leads to this very result. This is most important, for it brings the definition of an atom in a molecule into the realm of physics – the physics of an open system.

Equation (10) is transformed into a statement of physics by two rather remarkable consequences of the properties of the Laplacian of the electron density, $\nabla^2\rho(\mathbf{r})$. The first is that when Schrödinger's equation is satisfied, the integrand $\mathrm{f}(\psi, \nabla\psi)$ of Schrödinger's energy integral reduces to $-\mathrm{L}(\mathbf{r})$, the term proportional to $\nabla^2\rho(\mathbf{r})$ defined in Eq. (7). This same property obtains for the many-electron case, and for the Lagrangian density of the action integral in the general time-dependent case, persisting even in the presence of an electromagnetic field. So the term involving the variation of the surface may be re-expressed as $(\hbar^2/4\mathrm{m})\delta\mathrm{S}(\Omega; \mathbf{r}_s)\nabla^2\rho(\mathbf{r})$. The second of the remarkable consequences of the Lap-

lacian is its appearance in the constraint that determines the surface of the open system. This constraint is presented in detail in several places [3, 16], and is outlined only briefly here. Imposing the zero-flux surface condition expressed in Eq. (9) at every stage of the variation is equivalent to requiring that the variation of the integral of $\nabla^2\rho(\mathbf{r})$ over the region Ω vanish. This condition in turn enables one to replace the term $(\hbar^2/4\mathrm{m})\delta\mathrm{S}(\Omega;\mathbf{r}_\mathrm{s})\nabla^2\rho(\mathbf{r})$ with the term $-(\hbar^2/4\mathrm{m})\{(\nabla\psi^*)\delta\psi + \psi^*\delta\nabla\psi\} \bullet \mathbf{n}(\mathbf{r})$, a step outlined in my book (pp. 158 ff). Combining this result with the first of the surface terms in Eq. (10) results in a most remarkable transformation of the surface contribution, from one without any discernable physical content for a surface of unspecified form into an integral describing the flux in the variation of the quantum mechanical current density $\mathbf{j}(\mathbf{r})$ through a surface of zero-flux in $\nabla\rho$, as shown in Eq. (11) for a zero-flux surface.

$$\delta\mathscr{G}(\psi,\nabla\psi;\Omega) = -(\mathrm{i}\hbar/2)\oint \mathrm{dS}(\Omega;\mathbf{r}_\mathrm{s})\delta\mathbf{j}(\mathbf{r}) \bullet \mathbf{n}(\mathbf{r}) + \mathrm{cc} \tag{11}$$

The variation in $\mathbf{j}(\mathbf{r})$ is caused by variations in the wave function, as shown in Eq. (12):

$$\delta\mathbf{j}(\mathbf{r}) = (\hbar/2\mathrm{mi})\{\psi^*\nabla(\delta\psi) - (\nabla\psi^*)\delta\psi\} \tag{12}$$

Thus imposition of the *zero-flux boundary condition* on the variation of Schrödinger's functional causes it to be "stationary", requiring its variation to equal the surface flux in the current generated by the variations in ψ. Although unknown to us at the time of its derivation, Eq. (11) is a result of Schwinger's principle of stationary action for a time-independent system.

2.3.4
Getting Chemistry from $\delta\mathbf{G}(\psi,\nabla\psi;\Omega)$

Whereas Eq. (11), the expression for the variation in $\mathscr{G}(\psi,\nabla\psi;\Omega)$, links the definition of an atom with physics, it is not in a useable form. How does one describe the variation in the wave function that causes the change in $\mathbf{j}(\mathbf{r})$, Eq. (12)? Transforming Eq. (11) into an operational form is accomplished by replacing the variations in ψ with the action of quantum mechanical operators on ψ. All physical changes in the properties of a system, can be described by the action of the appropriate quantum mechanical operator $\hat{\mathbf{G}}$ on ψ. Thus, one makes the substitution $\delta\psi = -\varepsilon(\mathrm{i}/\hbar)\hat{\mathbf{G}}\psi$ where ε denotes an infinitesimal change resulting from the action of the operator $\hat{\mathbf{G}}$ on ψ, with $(\mathrm{i}/\hbar)$ present for dimensional reasons. $\hat{\mathbf{G}}(\mathbf{r})$ may be any linear Hermitian operator constructed from the electronic position and/or momentum coordinates of a single electron and is referred to as the generator of the change in the system. Its application to a wave function causes an infinitesimal unitary transformation in ψ. As explained below, it is this step that

completes the linking of the derivation of an atom in a molecule with Schwinger's principle of stationary action. With this substitution, Eq. (11) becomes:

$$\delta\mathscr{G}(\psi, \nabla\psi; \Omega) = -(\varepsilon/2)\left\{\oint \mathrm{dS}(\Omega; \mathbf{r}_s)\mathbf{j}_G(\mathbf{r}) \bullet \mathbf{n}(\mathbf{r}) + \mathrm{cc}\right\} \tag{13}$$

where $\mathbf{j}_G(\mathbf{r})$ is the current density for the property determined by the generator $\hat{\mathbf{G}}$:

$$\mathbf{j}_G(\mathbf{r}) = (\hbar/2\mathrm{mi})\{\psi^*\nabla(\hat{\mathbf{G}}\psi) - (\nabla\psi^*)(\hat{\mathbf{G}}\psi)\} \tag{14}$$

Comparison of Eqs. (12) and (14) shows that the variation in ψ appearing in $\delta\mathbf{j}(\mathbf{r})$ is replaced by the action of $\hat{\mathbf{G}}$ on ψ in the expression for $\mathbf{j}_G(\mathbf{r})$. Thus the variation in Schrödinger's energy functional for an open system bounded by a zero-flux surface – an atom in a molecule – is proportional to the surface flux in the current density of the infinitesimal generator causing the change in the system. Fluxes in currents normal to the bounding surface are the features that distinguish the physics of an open system from that of a total system, for which all surface terms vanish. There are, of course, no current flows in a stationary state in the absence of a magnetic field but, following Feynman: "We generalize the word "flux" to mean the surface integral of the normal component of a vector. Although it is not the flow of anything, we still call it the "flux"" [12]. A recent publication [17] claims that the surface flux term in Eq. (11), and hence in the general case, Eq. (13), vanishes for a stationary state, appearing to confuse the absence of a current in a stationary state with its "flux" through a surface, a claim that invalidates Schwinger's principle and the physics of an open system, and Eq. (16) that follows directly for Schrödinger's equation for a stationary state. A response to that paper has recently appeared [18].

The final, encompassing statement of the principle of stationary action for an open system in a stationary state is obtained by use of the equation of motion for the generator $\hat{\mathbf{G}}$. In the general time-dependent case, the time derivative of the average value of $\hat{\mathbf{G}}$, $\mathrm{d}\langle\hat{\mathbf{G}}\rangle/\mathrm{dt}$ is determined by the average of the commutator $(\mathrm{i}/\hbar)[\hat{\mathbf{H}}, \hat{\mathbf{G}}]$. This average vanishes for a molecule in a stationary state:

$$\langle\psi, [\hat{\mathbf{H}}, \hat{\mathbf{G}}]\psi\rangle = 0 \tag{15}$$

because of the Hermiticity of $\hat{\mathbf{H}}$, a property that does not usually apply to an open system [3]. The same commutator average does not however, vanish for an atom in a molecule, the contribution from the commutator being balanced by the surface flux in the current of $\hat{\mathbf{G}}$, a result readily obtained from Schrödinger's equation [3] and given in Eq. (16):

$$(\mathrm{i}/\hbar)\langle\psi, [\hat{\mathbf{H}}, \hat{\mathbf{G}}]\psi\rangle_\Omega + \mathrm{cc} = \oint \mathrm{dS}(\Omega; \mathbf{r}_s)\mathbf{j}_G(\mathbf{r}) \bullet \mathbf{n}(\mathbf{r}) + \mathrm{cc} \tag{16}$$

Equation (16) yields the same surface term obtained in the variation of $\mathcal{G}(\psi, \nabla\psi; \Omega)$ and substitution of this result into Eq. (13) yields the atomic statement of the principle of stationary action for a stationary state:

$$\delta\mathcal{G}(\psi, \nabla\psi; \Omega) = -(\varepsilon/2)\{(\mathrm{i}/\hbar)\langle\psi, [\hat{\mathbf{H}}, \hat{\mathbf{G}}]\psi\rangle_\Omega + \mathrm{cc}\} \tag{17}$$

This statement determines the physics of an atom in a molecule [3]. All of the theorems of quantum mechanics are obtained by appropriate choice of the generator $\hat{\mathbf{G}}$ and thus it applies to every property determined by a quantum mechanical operator – to all of the measurable properties of a system. The atomic statement of the virial theorem for example, is obtained by setting $\hat{\mathbf{G}} = \hat{\mathbf{r}} \bullet \hat{\mathbf{p}}$, the product of the electronic position and momentum coordinates. *Equation (17) applies to any system bounded by a zero-flux surface and thus a single principle provides the quantum mechanical description of the total system and of its constituent atoms.* Indeed one may regard the physics of some total system – of the entire molecule – as a special limiting case of the more general expression pertaining to an open system given in Eq. (17). Thus when Ω refers to the total molecule, the commutator average equals zero, Eq. (15), and the variation in Schrödinger's functional becomes stationary in the usual sense that $\delta\mathcal{G}(\psi, \nabla\psi; \Omega) = \delta\mathcal{G}(\psi, \nabla\psi) = 0$. It is important to note that the derivation of Eq. (17) yields Schrödinger's equation and all of the theorems of quantum mechanics. Thus a single principle serves to completely determine the physics of a stationary state, of the total system, and of its constituent atoms, all as a consequence of Schwinger's principle [3].

2.4 From Schrödinger to Schwinger

2.4.1 From Dirac to Feynman and Schwinger

The generalization of the variation of Schrödinger's energy functional to obtain the physics of an open system, Eq. (17), necessitated the generation and retention of terms resulting from the variation of the surface and of the variation of the wave function on the surface. Although such terms are discarded in the principle of least action when applied to a total system, they play a crucial role in the physics of an open system. It is this step that transforms Schrödinger's approach into Schwinger's principle of stationary action. We can only briefly hint at the beauty of the chain of reasoning that leads from Dirac to Feynman and Schwinger and ultimately to an atom in molecule. Dirac introduced transformation theory into quantum mechanics. This is the underlying mathematical formulation of the new physics which consists of the general mathematical scheme of linear operators and state vectors with its associated probability interpretation. In doing so, he stressed how the theory of infinitesimal unitary transformations in

quantum mechanics parallels the infinitesimal canonical transformations of classical theory thereby "providing the mathematical foundation of the analogy between classical and quantum equations of motion". Thus identification of the variations in ψ with the action of infinitesimal generators $\varepsilon\hat{\mathbf{G}}$ to obtain Eq. (17) has deep implications beyond yielding operational expressions for the mechanics of an open system. In 1933 such thinking led Dirac to write a paper [19] wherein he posed the question of what would correspond to the limiting classical expression for the quantum transition amplitude that determines the dynamic behavior of the system with time. Dirac was asking for the correspondence of quantum mechanics with the Lagrangian method of classical mechanics, a formulation he considered to be more fundamental than that based on Hamiltonian theory. Dirac proposed that the transition amplitude be given by an exponential of $(i/\hbar)\mathscr{W}_{12}$, where $\mathscr{W}_{12}$ is the classical action evaluated along the unique path that causes it to be stationary, a proposal that set the stage for the reformulations of quantum mechanics proposed by Feynman in his path integral approach [8] and by Schwinger in his quantum dynamical principle [4]. Both approaches enable one to derive the commutation relations, as opposed to the need to postulate them in the Hamiltonian approach [20].

Schwinger realized that by retaining the variations on the boundary of the space–time volume of the action integral swept out by the temporal evolution of a system, followed by their identification with the generators of infinitesimal unitary transformations, he could combine the action principle which yields Schrödinger's equation of motion with Dirac's transformation theory and thus recover "all of physics" from a single dynamical principle – the principle of stationary action [4]. The retention of the variations both of and on the zero-flux surface at a single time t followed by their identification with the action of infinitesimal generators is precisely the step made in the generalization of Schrödinger's variation of his energy functional leading to Eq. (17). These variations give rise to the same surface terms that are found in Schwinger's formulation and are dealt with in precisely the same manner. While we proceeded out of necessity, Schwinger did so by choice.

2.4.2
From Schwinger to an Atom in a Molecule

In 1978 it was realized that the generalization of the variation of Schrödinger's functional was but a special case of the general statement of physics provided by Schwinger's principle of stationary action and the theory of an atom in a molecule was readily extended to the general time-dependent case [16, 21]. Criticisms of the extension have been responded to in full [22, 23]. The operational statement of Schwinger's principle of stationary action is given in terms of the variation of the Lagrangian, Eq. (18):

$$\delta\mathscr{L}(\Psi, \nabla\Psi, \mathrm{t}; \Omega) = (\varepsilon/2)\{(\mathrm{i}/\hbar)\langle\Psi, [\hat{\mathbf{H}}, \hat{\mathbf{G}}]\Psi\rangle_{\Omega} + \mathrm{cc}\} \tag{18}$$

The form of this expression is to be compared with that obtained from variation of Schrödinger's energy functional for a stationary state, Eq. (17), from which it differs only by the expected change in sign. On evaluation of the variation in the Lagrangian, Eq. (18) yields Eq. (19) for the equation of motion for a generator $\hat{\mathbf{G}}(\mathbf{r})$, a function of the coordinate $\mathbf{r}$ of a single electron [24].

$$\begin{aligned}(1/2)\mathrm{N}\int_{\Omega}\mathrm{d}\mathbf{r}\int\mathrm{d}\mathbf{r}'\partial\{\Psi^{*}\hat{\mathbf{G}}(\mathbf{r})\Psi+\mathrm{cc}\}/\partial\mathrm{t}\\ =(1/2)\{(\mathrm{i}/\hbar)\langle\Psi|[\hat{\mathrm{H}},\hat{\mathbf{G}}(\mathbf{r})]|\Psi\rangle_{\Omega}+\mathrm{cc}\}\\ -(1/2)\oint\mathrm{d}\mathbf{S}(\mathbf{r}_{\mathrm{s}};\Omega)\bullet\{\mathbf{J}_{\mathrm{G}}(\mathbf{r}_{\mathrm{s}})+\mathrm{cc}\}\end{aligned}\tag{19}$$

The most recent derivation of this result is given in terms of an augmented Lagrange-function operator, termed a *proper operator*, because only its variation recovers the equations of motion predicted by Schrödinger's field equation [25]. The resulting open systems are termed *proper open systems*.

The term $\mathrm{N}\int\mathrm{d}\mathbf{r}'\{\Psi^{*}\hat{\mathbf{G}}(\mathbf{r})\Psi+\mathrm{cc}\}$, whose time derivative appears in Eq. (19), defines the density of the property G associated with the operator $\hat{\mathbf{G}}(\mathbf{r})$. The integration symbol denotes that the operator is averaged over the coordinates of all the electrons except those denoting the position $\mathbf{r}$, the coordinate of the electron to be integrated over Ω. All open-system properties, including the energy, are thus defined in terms of a real-space density distribution, in the same manner as is the electron density where the integrand is simply $\Psi^{*}\Psi$, with the result that the value of some property for the total system equals the sum of its atomic contributions. The symbol $\langle\ \rangle_{\Omega}$ implies the same averaging. The LHS vanishes for a stationary state as does the surface term for a total isolated system. The variational result given in Eq. (19) determines the physics of *all measurable properties*, the properties of a total system being a special limiting case of those for an open system.

QTAIM is predicated on the premise that a theory of an atom in a molecule must predict what can be measured in the laboratory. It is this philosophy that underlies Hans Bethe's view of science: ''its great advantage is you can prove something is true or something is false'', a statement he further paraphrased as ''In science, you know you know'' [26]. Therefore, the single necessary and sufficient criterion for determining the relevance of QTAIM atoms to chemistry is agreement of the predicted additive atomic and group contributions with their experimental values, agreement with observation being the only test of theory. QTAIM has demonstrably satisfied this experimental criterion, with examples from heats of formation, electric and magnetic susceptibilities, and molar volumes, and from the recent demonstration of the recovery of the measured atomic contributions to infrared intensities obtained from the atomic polar tensor [27, 28]. This agreement is but a necessary consequence of quantum mechanics predicting all measurable properties, Eq. (19). A recent paper provides a comprehensive summary of the recovery of measured properties by QTAIM [29].

The interested reader is referred to the original papers and my book for a complete description of the extension of Schwinger's principle of stationary action to

an open system. There are numerous important nuances that one must take into account before one attempts a critique. An important example of this stems from the fundamental step in Schwinger's approach – the variation of the state vector on the space–time boundary of the open system with the variation of the surface, followed by identification of these variations with the generators of infinitesimal unitary transformations. Thus the theory requires the use of a special class of trial functions whose variation will correspond to continuous changes in the dynamic variables of the physical system. This is the very requirement needed to ensure the applicability of the zero-flux surface condition as the defining constraint of a proper open system.

Equation (19) is the bridge that links the Lagrangian and Heisenberg representations of quantum mechanics thereby linking theory with experiment. Table 2.1

Table 2.1 The atomic theorems for molecules and crystals obtained for several important generators.

Atomic force theorem $\hat{\mathbf{G}} = \hat{\mathbf{p}}$

$$\mathrm{m}\int_{\Omega} \mathrm{d}\mathbf{r}\partial\mathbf{J}(\mathbf{r})/\partial \mathrm{t} = \int_{\Omega} \mathrm{d}\mathbf{r}\int \mathrm{d}\tau' \Psi^{*}(-\nabla\hat{\mathrm{V}})\Psi + \oint \mathrm{dS}(\mathbf{r})\sigma(\mathbf{r})\cdot\mathbf{n}(\mathbf{r})$$

Atomic virial theorem $\hat{\mathbf{G}} = \hat{\mathbf{r}}\cdot\hat{\mathbf{p}}$

$$\mathrm{m}\int_{\Omega} \mathrm{d}\mathbf{r}\mathbf{r}\cdot\partial\mathbf{J}(\mathbf{r})/\partial \mathrm{t} = 2\mathrm{T}(\Omega) + \int_{\Omega} \mathrm{d}\mathbf{r}\int \mathrm{d}\tau' \Psi^{*}(-\mathbf{r}\cdot\nabla\hat{\mathrm{V}})\Psi + \oint \mathrm{dS}(\mathbf{r})\mathbf{r}\cdot\sigma(\mathbf{r})\cdot\mathbf{n}(\mathbf{r})$$

Atomic torque theorem $\hat{\mathbf{G}} = \hat{\mathbf{r}}\times\hat{\mathbf{p}}$

$$\mathrm{m}\int_{\Omega} \mathrm{d}\mathbf{r}\mathbf{r}\times\partial\mathbf{J}(\mathbf{r})/\partial \mathrm{t} = \int_{\Omega} \mathrm{d}\mathbf{r}\int \mathrm{d}\tau' \Psi^{*}(-\mathbf{r}\times\nabla\mathrm{V})\Psi - \oint \mathrm{dS}\sigma(\mathbf{r})\times\mathbf{r}\cdot\mathbf{n}$$

Atomic current theorem $\hat{\mathbf{G}} = \hat{\mathbf{r}}$

$$\int_{\Omega} \mathrm{d}\mathbf{r}\mathbf{r}\partial\rho(\mathbf{r})/\partial \mathrm{t} = \int_{\Omega} \mathrm{d}\mathbf{r}\,\mathbf{J}(\mathbf{r}) - \oint \mathrm{dS}\mathbf{n}(\mathbf{r})\cdot\mathbf{J}(\mathbf{r})\mathbf{r}$$

Atomic continuity theorem $\hat{\mathbf{G}} = \hat{\mathrm{N}}$

$$\int_{\Omega} \mathrm{d}\mathbf{r}\partial\rho(\mathbf{r})/\partial \mathrm{t} = -\oint \mathrm{dS}\mathbf{J}(\mathbf{r})\cdot\mathbf{n}(\mathbf{r})$$

Atomic power theorem $\hat{\mathbf{G}} = \hat{\mathbf{p}}^2/2\mathrm{m}$, written without $1/2\mathrm{m}$

$$\int_{\Omega} \mathrm{d}\mathbf{r}\partial\rho_{\mathbf{p}^2}(\mathbf{r})/\partial \mathrm{t} = \int_{\Omega} \mathrm{d}\mathbf{r}\int \mathrm{d}\tau'(\hbar/i)\{(\Psi\nabla\Psi^{*} - \Psi^{*}\nabla\Psi)\cdot\nabla\hat{\mathrm{V}}\} + \oint \mathrm{dS}\,\mathrm{Re}\{\mathbf{J}_{\mathbf{p}^2}(\mathbf{r})\}$$

Quantum stress tensor density

$$\sigma(\mathbf{r}) = (\hbar^2/4\mathrm{m})\int \mathrm{d}\tau'\{(\nabla\nabla\Psi^{*})\Psi - \nabla\Psi^{*}\nabla\Psi - \nabla\Psi\nabla\Psi^{*} + \Psi^{*}\nabla\nabla\Psi\}$$

Quantum vector current density

$$\mathbf{J}(\mathbf{r}) = (\hbar/2\mathrm{m}i)\int \mathrm{d}\tau'\{\Psi^{*}\nabla\Psi - \Psi\nabla\Psi^{*}\}$$

lists the atomic theorems for molecules and crystals obtained for several important generators [30]. One notes the presence in these theorems of the electron density $\rho(\mathbf{r})$, the current density $\mathbf{j}(\mathbf{r})$ and the quantum stress tensor $\ddot{\sigma}(\mathbf{r})$, quantities introduced by Schrödinger in 1926 and held by him to be essential to understanding the mechanical, electrical, and magnetic properties of matter. For example, the atomic virial theorem is required for definition of the energy of an open system and of the pressure acting on its surface. The atomic current theorem is required to determine the atomic contribution to the molecular diamagnetic susceptibility arising from the flux in the position-weighted flux in the induced current through the atomic surface. The use of more of these theorems will be illustrated in the contributions to this book.

2.5 Molecular Structure and Structural Stability

2.5.1 Definition of Molecular Structure

The remarkable encompassing physical aspect of the theory is that the same topology that defines an atom in a molecule, leads to a theory of molecular structure and structural stability. The response of the electron density to the interaction between two atoms is ubiquitous, resulting in the formation of a $(3,-1)$ critical point whose associated trajectories define not only the presence of the zero-flux interatomic surface but also delineate a line of maximum electron density that links the nuclei of neighboring atoms – the "bond path" [3, 31–33]. The network of bond paths generate a *molecular graph* that defines a system's structure. The topological structures have been shown to recover the "chemical structures" in a multitude of systems, in terms of densities obtained from both theory and experiment, structures that were previously inferred from classical models of bonding in conjunction with observed physical and chemical properties [34].

A bond path meets all the physical requirements set by the Ehrenfest, Feynman and virial theorems that the atoms be bonded to one another [34]; the two atoms experience an attractive Ehrenfest force drawing their atomic basins together – no Feynman force, either attractive or repulsive, acts on the nuclei, because of the balancing of the repulsive and attractive forces by accumulation of electron density in the binding region. This same accumulation leads to a reduction of the electron–nuclear potential energy, the magnitude of which exceeds the increases in the electron and nuclear repulsion energies, resulting in a decrease in the potential energy equal to twice the decrease in the total energy, all as demanded by the virial theorem. Thus a bond path is indicative of the accumulation of density between the nuclei that is necessary for the presence of attractive Ehrenfest forces, for balancing of the Feynman forces on the nuclei, and for the decrease

in energy. Its presence is both necessary and sufficient for two atoms to be bonded to one another [35, 36].

2.5.2
Prediction of Structural Stability

A central concept in the analysis of the stability of gradient vector field is the equivalence relationship. Two vector fields are said to be equivalent if every trajectory of one field can be mapped on to a corresponding trajectory of the other. By application of this equivalence relationship to the gradient vector field of the electron density in *behavior space*, one arrives at a partitioning of nuclear configuration space R^Q, the *control space*, into a finite number of disjoint regions, the structural regions, each of which is characterized by a unique molecular graph. The structural regions form a dense open subset of R^Q and a point belonging to such a region is a stable structure and is called a *regular point*. The catastrophe set C is the collection of all structurally unstable points in R^Q and serves as the union of the boundaries of all the structural regions. The result is a *structure diagram, a diagram that defines all possible structures and all structural changes linking the structures for a given system* [3, 37]. The Palis–Smale theorem of structural stability [38] shows that a change in structure can occur by only one of two possible mechanisms – the bifurcation mechanism arising from the formation of a degenerate critical point in the density or through the conflict mechanism wherein the manifolds of two critical points intersect in what is a manifestly unstable manner.

The topological theory of molecular structure and structural stability leads to several important observations [3]. A molecular geometry, a point in R^Q, should be distinguished from a molecular structure, which represents an open region of R^Q, that is, structure is generic. Motion in R^Q changes the geometry, but leaves the structure unchanged for motion within the open region associated with a given structural region. A change in structure is an abrupt and discontinuous process and occurs when a system point crosses a boundary at a nuclear configuration in R^Q belonging to the catastrophe set separating two structural regions. One finds these ideas being increasingly applied to a wide range of problems.

2.6
Reflections and the Future

2.6.1
Reflections

It is indeed a pleasure to write this chapter for a book that demonstrates the remarkable progress that has been made in the development and application of QTAIM. All problems at the atomic level are subject to study by QTAIM and the

scope of its application is forever widening. The use of the atomic theorems derived from the Heisenberg equations of motion Eq. (19), some of which are displayed in Table 2.1 [30], extends beyond the asking of chemical questions of bonding, structure, and reactivity. The theorems apply to all questions about the behavior of matter at the atomic level. Indeed, there are many problems that specifically require the physics of an open system for their statement and solution. Examples are the operation of the atomic force and electron tunneling microscopes [39], defects in solids [40], the quantum definition of pressure [41], and the polarization of a dielectric [42]. The atomic statement of the Ehrenfest force is the equation of motion for an open system. It is capable of describing the motion of an adsorbed atom on the surface of a substrate or of the forces required for manipulation of individual atoms or molecules, thereby providing a basis for nanotechnology. The physical understanding obtainable from the use of the theorems for an open system is only beginning to be explored.

Molecular orbital theory is *the* theory for the understanding and prediction of the electronic structure of molecules, predicting the ordering and classification of many-electron states in terms of one-electron states, and it is indispensable to all chemists for the understanding of the properties of many-electron systems. Orbital ordering and state classification forms the basis for the application of the ''second-order Jahn–Teller'' symmetry rule [43] that underlies Fukui's frontier orbital theory [44] and orbital conservation [45].

Unlike QTAIM, which builds upon the chemical concept of a functional group with characteristic properties, molecular orbital theory, as Libit and Hoffmann point out [46], is incapable of recovering this concept, because each molecular orbital extends over the entire molecule. Recent papers illustrate the complementary roles of molecular orbital theory and QTAIM and how QTAIM provides the possibility of assessing the viability of orbital models [47, 48].

To argue, as some do, that QTAIM overlooks explanations of bonding afforded by simple orbital models is wrong. What QTAIM does do is enable one to go beyond the models. To state that QTAIM, by finding a bond path in Ar_2 for example, fails to predict the absence of electron pair bonding between closed-shell molecules is based on improper use of the orbital model [49]. The orbital model when properly applied and understood using QTAIM, predicts weak bonding between closed shell systems. It is a travesty to claim that molecular orbital theory cannot account for the bonding between closed-shell systems by insisting that one terminates the theory at the single determinant level. Electron correlation is known to be responsible for the bonding in such cases and is readily accounted for by inclusion of the interaction of excited configurations with the ground state. Thus, for example, a CI calculation predicts both bonding *and* the presence of a bond path in rare gas dimers [50], a result in accord with the experimental detection of bound He_2 [36].

In summary, there is no conflict of QTAIM with molecular orbital theory. The conflict is with those who introduce nonphysical concepts, postulating the presence of repulsive forces in systems wherein no definable repulsive forces act on the density or on the nuclei [49] (appropriate responses being given in [51, 52]).

One can hope that younger scientists, not having been exposed to models that have outlived any usefulness they might have once enjoyed, will place their trust in physics.

2.6.2 The Future

A necessary next step in the development of the theory is its extension beyond the fixed nucleus approximation of the Born–Oppenheimer procedure. The topology of the electron density is a consequence of, and summarizes the physics that underlies, the form of matter. Whatever new topological features the charge density may be found to exhibit as a consequence of its averaging over nuclear motions, they may be incorporated into an expanded theory to provide a still deeper understanding of the behavior of matter at the atomic level.

Relativistic effects cause no problem, as the mathematical formalism of Schwinger's and Feynman's approaches is manifestly co-variant with regard to Lorentz transformations, if one adheres to a relativistically invariant Lagrangian [7], as is done in the development of QTAIM.

James Anderson, a student of Dr. Paul Ayers, has recently extended QTAIM to the relativistic domain using the ZORA Hamiltonian, as a result of attending my graduate course. Schwinger's theory is, of course, relativistically invariant, but there are a number of crucial steps involved in its extension to an open system, not the least of which is the zero-flux boundary condition. He finds QTAIM to be 'robust', the entire theory comes through unchanged and can be applied with the same zero-flux boundary condition across the periodic table – including the actinides. It is beautiful the way in which all of the important properties of the relativistic Lagrangian mimic the essential properties of the non-relativistic case, all of the derived relativistic expressions reducing to their non-relativistic forms in the limit of infinite c. I can see no reason for anyone doubting that the zero-flux boundary is a fundamental property of matter, providing the basis for the generalization of physics to its atomic constituents.

In many problems, the system of interest is an open system embedded in a much larger one. For example, in the development of molecular devices in nanotechnology one wishes to determine the conductivity of an organic molecule linking two conductors. In biological systems one's interests may focus on just the active or binding site of an enzyme. Can one define and study the open system of interest, rather than resorting to existing "embedding" methods. The properties of an open system are totally determined by its bounding surface [53]. The determination of the open system in a simpler environment, then its transfer to the system of interest, is a possibility that has already been successfully explored in the construction of biological molecules, in which the groups of interest have high transferability and close matching of the interatomic surfaces is possible.

The open system variation principle, Eq. (11), offers the possibility of obtaining the wave function for the entire system by performing the variation over just the open system of interest, including its surface, and requiring that the variations

equal the surface flux in the infinitesimal current generated by the variations in ψ. The variation is, of course, subject to the constraint that at every stage the system be bounded by a zero-flux surface. Baranger [54] has shown that at Hartree–Fock, minimization of the energy (of the total system) is equivalent to satisfaction of the vanishing of $\langle\psi[\hat{H},\hat{G}]\psi\rangle$, Eq. (15), for all one-body operators $\hat{G}$, and has given a procedure for doing so. One could apply the same procedure to a open system by noting that satisfaction of Eq. (11) is equivalent to satisfying Eq. (16) for a proper open system, and thus requires that the commutator average over the open system be given by the surface flux in the corresponding currents for all one-body operators $\hat{G}$.

What is, to me, among the most important of the results derived from the development of the theory of atoms in molecules is the paralleling behavior in the form and properties found for a proper open system. This observation is valid at all levels of transferability of the density – from near transferability, as found for the Li atom in its hydride, oxide, and fluoride molecules in the original 1972 paper (the observation that sparked the development of the theory) to the essentially perfect transferability observed, for example, for the transferable methyl and methylene groups in linear hydrocarbons or amino acid residues in a polypeptide [55]. Although it is an obvious physical necessity that form determine properties, it is only through the atoms of QTAIM that this condition is realized, and it is striking. One need only view the previously illustrated [56] remarkable degree of paralleling transferability of the electron density, the kinetic energy density, and the virial field (the potential energy density) of the methylene groups in butane and pentane as an example. The virial field is a real-space representation of the average effective potential experienced by a single electron in a many-electron system. It is the most short-range description possible of this interaction potential [57] and its integral over all space yields the total potential energy of the molecule in the fixed nucleus approximation [58]. Add to this the observation that it is structurally homeomorphic with the electron density [59] and it would seem to be a promising starting point for investigation of the energy $\leftrightarrow$ density relationship. The local statement of the atomic virial theorem, by relating the Laplacian of the density to the kinetic energy density and the virial field, provides a link between a property of the density with energy [3, 16].

It is clear from the tabulations of experimentally derived group properties that group additivity must often be only apparent, as is found to apply when a group is unavoidably perturbed by its environment. The additivity in these cases is a result of *compensatory transferability*, the change in a property, energy for example, and in the charge that is experienced by one group being equal and opposite to the changes experienced by the other group [60]. Charge is, of course, necessarily conserved, but for this conservation to be paralleled by all properties is remarkable, one that can be illustrated by the data in the tables of Benson et al. [61, 62] and in theoretical calculations [53, 60]. There seems to be a Le Chatellier principle at work – one that states that two open systems brought into contact respond in such a way as to minimize the overall changes in their form and properties [63]. Is it possible to formulate an extremization principle that minimizes the sum of the energy changes of the two open system when brought into contact?

Chemistry is a consequence of the short-range nature of the one-electron density matrix that determines all the mechanical properties of an atom in a molecule [57] with the additional important proviso that all of the necessary physical information is obtained in its expansion up to second-order with regard to both the diagonal and off-diagonal terms [3]. The diagonal terms yield the density $\rho(\mathbf{r})$, its gradient vector field $\nabla\rho(\mathbf{r})$, and its dyadic $\nabla\nabla\rho(\mathbf{r})$ whose trace yields the Laplacian $\nabla^2\rho(\mathbf{r})$ and determines the critical points in $\nabla\rho(\mathbf{r})$ and hence in $\rho(\mathbf{r})$. The off-diagonal terms yield the current density $\mathbf{j}(\mathbf{r})$, the stress tensor $\overleftrightarrow{\sigma}(\mathbf{r})$ and the tensor $\nabla\mathbf{j}(\mathbf{r})$ whose properties determine the critical points in $\mathbf{j}(\mathbf{r})$ [64]. There is much to be studied. Whereas the topologies of $\rho(\mathbf{r})$ [37] and of $\mathbf{j}(\mathbf{r})$ [64] have been completely characterized and related to the physical properties of the system, the same is not true of the stress tensor $\overleftrightarrow{\sigma}(\mathbf{r})$. The topology of $\overleftrightarrow{\sigma}(\mathbf{r})$, whose properties determine the local mechanics of the density, its trace equaling the virial field and the local statement of the virial theorem, is largely unstudied. Its eigenvalues and eigenvectors at a degenerate critical point in the density that is indicative of a change in structure could summarize the mechanical consequences. Similarly, the topology of the Ehrenfest force field, a vector field defined by $\nabla \bullet \overleftrightarrow{\sigma}(\mathbf{r})$, could prove invaluable in understanding the mechanics of the density. A surface of zero-flux in this field would demark a region of space for which the Ehrenfest force vanishes. The final relationship between the density and the energy will have to account for their paralleling behavior as evinced by the atoms of QTAIM, so it seems reasonable that the observations of QTAIM regarding the role of the density in the determination of atomic properties could serve as a starting point in the search for such a relationship [57, 58].

References and Notes

1 R. F. W. Bader, P. M. Beddall, *J. Chem. Phys.* **1972**, *56* 3320–3329.
2 S. Srebrenik, R. F. W. Bader, *J. Chem. Phys.* **1975**, *63* 3945–3961.
3 R. F. W. Bader, *Atoms in molecules: a quantum theory*, Oxford University Press, Oxford UK **1990**.
4 J. Schwinger, *Phys. Rev.* **1951**, *82* 914–927.
5 J. C. Slater, *J. Chem. Phys.* **1972**, *57* 2389.
6 E. Schrödinger, *Ann. D. Phys.* **1926**, *79* 361.
7 W. Yourgrau, S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory*, Dover Pub. Inc., New York **1979**.
8 R. P. Feynman, *Rev. Mod. Phys.* **1948**, *20* 367–387.
9 M. Gell-Mann, *Physics To-Day* **1989**, *42* 50.
10 Y & M discuss the unbounded enthusiasm that Planck held for the principle of least action, dubbing it "the most comprehensive of all physical laws which governs equally mechanics and electrodynamics."
11 R. F. W. Bader, *Monatshefte für Chemie* **2005**, *136* 819–854.
12 R. P. Feynman, R. B. Leighton, M. Sands, *The Feynman Lectures on Physics. Vol. II*, Addison–Wesley Publishing Co. Inc., Reading, MA **1964**.
13 One should distinguish between the use the calculus of variations to obtain *Schrödinger's equation* from that of the more familiar "variational principle", for obtaining approximate solutions to his equation. The variational principle states that the energy obtained by averaging the

Hamiltonian operator over a trial function, which usually includes variational parameters in the minimization of the "variational integral", is an upper bound to the true energy E.

14 The Lagrangian approach is used throughout mechanics and in every case, from diffusion to electromagnetics to general relativity to Boltzmann's law, to cite just a few examples, the expression for the Lagrangian contains a squared gradient term of the form representing the kinetic energy as in Eq. (4). Frieden (Frieden, R. B. *Physics from Fisher Information*, Cambridge University Press, 1998) has argued that the presence of such a term is natural for all fields and provides a measure of the field's intrinsic information called the "Fisher information".

15 R. F. W. Bader, H. J. T. Preston, *Int. J. Quantum Chem.* **1969**, *3* 327–347.

16 R. F. W. Bader, T. T. Nguyen-Dang, *Ad. Quantum Chem.* **1981**, *14* 63–124.

17 A. M. Pendás, M. A. Blanco, E. Francisco, *Chem. Phys. Lett.* **2006**, *417* 16–21.

18 R. F. W. Bader, *Chem. Phys. Lett.* **2006**, *426*, 226–228.

19 P. A. M. Dirac, *Physik. Zeits. Sowjetunion* **1933**, *3* 64.

20 An account of Dirac's proposal using the multiplicative law of transformation functions and its consequences for Feynman's path integral is given in Section 8.6 of my book. Schwinger's use of his differential characterization of the transformation function as the basis for his quantum action principle is given in Section 8.2.

21 R. F. W. Bader, S. Srebrenik, T. T. Nguyen Dang, *J. Chem. Phys.* **1978**, *68* 3680–3591.

22 R. F. W. Bader, *Theor. Chem. Acc.* **2001**, *105* 276–283.

23 R. F. W. Bader, *Theor. Chem. Acc.* **2002**, *107* 381–382.

24 The expression for the total time derivative of the expectation value for an open system contains in addition to the terms shown in Eq. (19), a surface integral of the change in the surface with time weighted by the property density of the observable G.

25 R. F. W. Bader, *Phys. Rev.* **1994**, *B 49* 13348–13356.

26 H. Bethe, *New York Times Obituary* **2005**, March 18.

27 R. L. A. Haiduke, R. E. Bruns, *J. Phys. Chem. A* **2005**, *109* 2680–2688.

28 P. H. César, S. H. D. Faria, J. V. da Silva, R. L. A. Haiduke, R. E. Bruns, *Chem. Phys.* **2005**, *317* 35–42.

29 C. F. Matta, R. F. W. Bader, *J. Phys. Chem. A* **2006**, *110* 6365–6371.

30 R. F. W. Bader, P. L. A. Popelier, *Int. J. Quantum Chem.* **1993**, *45* 189–207.

31 G. R. Runtz, R. F. W. Bader, R. R. Messer, *Can. J. Chem.* **1977**, *55* 3040–3045.

32 R. F. W. Bader, S. G. Anderson, A. J. Duke, *J. Am. Chem. Soc.* **1979**, *101* 1389–1395.

33 R. F. W. Bader, in *M. T. P. International Series of Science, Theoretical Chemistry, Vol. II*, (Eds.: C. A. Coulson, E. D. A. Buckingham), Butterworths, **1975**, pp. 43–79.

34 R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102* 7314–7323.

35 For a system in a stationary state, the wave function and the electron density it determines are such as to minimize the total energy. Thus the formation of an *atomic interaction line* – a line of maximum density linking a pair of nuclei – is associated with the minimization of the energy of interaction between the atoms it links. In the case where the attractive restoring forces return a system to its equilibrium geometry it is termed a *bond path*. All bond paths are transformed into atomic interaction lines when the separation between the atoms is compressed to the point where repulsive forces dominate the interaction, as determined by the kinetic energy T exceeding the negative of the total energy E, $T > -E$. One tires of reading the criticism of a bond path being present in the case of "repulsive interactions" such as He_2. At its equilibrium separation,

He_2 is bound and exhibits a bond path. Indeed, the bound He_2 molecule has been observed experimentally and reported in a paper entitled "The weakest bond: Experimental observation of helium dimer" [36]. Like all molecules, it becomes repulsive when compressed and the bond path is transformed into an atomic interaction line. It is equally important to delineate repulsive as well as bonded interactions.

36 F. Luo, G. C. McBane, G. Kim, C. F. Giese, R. Gentry, *J. Chem. Phys.* **1993**, *98* 3564–3567.

37 R. F. W. Bader, T. T. Nguyen-Dang, Y. Tal, *Rep. Prog. Phys.* **1981**, *44* 893–948.

38 J. Palis, S. Smale, *Pure Math.* **1970**, *14* 223.

39 R. F. W. Bader, *Phys. Rev. B* **2000**, *61* 7795–7802.

40 R. F. W. Bader, J. A. Platts, *J. Chem. Phys.* **1997**, *107* 8545–8553.

41 R. F. W. Bader, M. A. Austen, *J. Chem. Phys.* **1997**, *107* 4271–4285.

42 R. F. W. Bader, *Mol. Phys.* **2002**, *100* 3333–3344.

43 R. F. W. Bader, *Can. J. Chem.* **1962**, *40* 1164–1175. This paper presents an orbital correlation diagram from reactants to transition state to products that demonstrates how in an allowed reaction, a new low-lying excited state is generated as a consequence of the non-crossing rule for orbitals of the same symmetry, creating a transition density that facilitates the motion of the system over the energy barrier. It is a simple extension to the case of a forbidden reaction where such a state would not be generated, the differing symmetries of the pertinent orbitals failing to prevent their intersection, resulting in an increasing energy of the reactants.

44 K. Fukui, *Acc. Chem. Res.* **1971**, *4* 57–64.

45 R. B. Woodward, R. H. Hoffmann, *The conservation of orbital symmetry*, Verlag Chemie, Weinheim **1970**.

46 L. Libit, R. Hoffmann, *J. Am. Chem. Soc.* **974**, *96* 1370–1383.

47 F. Cortés-Guzmán, R. F. W. Bader, *Coordination Chem. Rev.* **2005**, *249* 633–662.

48 R. F. W. Bader, C. F. Matta, F. Cortés-Guzmán, *Organometallics* **2004**, *23* 6253–6263.

49 J. Poater, M. Solà, F. M. Bickelhaupt, *Chem. Eur. J.* **2006**, *12* 2902.

50 J. Hernández-Trujillo, R. F. W. Bader, *J. Phys. Chem. A* **2000**, *104* 1779–1794.

51 R. F. W. Bader, *Chem. Eur. J.* **2006**, *12* 2896–2901.

52 R. F. W. Bader, *Int. J. Quantum Chem.* **2003**, *94* 173–177.

53 R. F. W. Bader, F. J. Martin, *Can. J. Chem.* **1998**, *76* 284–291.

54 M. Baranger, W. A. Benjamin, **1963**.

55 R. F. W. Bader, C. F. Matta, F. J. Martin, *Quantum Medicinal Chemsitry, Eds. Carloni and Alber, F.*, Wiley–VCH, Weinheim, Germany **2003**.

56 R. F. W. Bader, *Can. J. Chem.* **1998**, *76* 973–988.

57 R. F. W. Bader, *Int. J. Quantum Chem.* **1995**, *56* 409–419.

58 R. F. W. Bader, in *The Fundamentals of Electron Density, Density Matrix and Density Functional Theory of Atoms, Molecules and the Solid State* Eds.: N. I. Gidopoulos, S. Wilson, Kluwer Academic Publishers, Netherlands **2003**, pp. 185–193.

59 T. A. Keith, R. F. W. Bader, Y. Aray, *Int. J. Quantum Chem.* **1996**, *57* 183–198.

60 R. F. W. Bader, D. Bayles, *J. Phys. Chem. A* **2000**, *104* 5579–5589.

61 S. W. Benson, J. H. Buss, *J. Chem. Phys.* **1958**, *29* 546–572.

62 S. W. Benson, F. R. Cruickshank, D. M. Golden, G. R. Haugen, H. E. O'Neal, A. S. Rodgers, R. Shaw, R. Walsh, *Chem. Rev.* **1969**, *69* 279.

63 F. Cortés-Guzmán, R. F. W. Bader, *J. Phys. Org. Chem.* **2004**, *17* 95–99.

64 T. A. Keith, R. F. W. Bader, *J. Chem. Phys.* **1993**, *99* 3669–3682.

3
Atomic Response Properties

Todd A. Keith

3.1
Introduction

This chapter describes some applications of the quantum theory of atoms in molecules (QTAIM) [1] with the objective of an atomic description of molecular response properties. For the purposes of this chapter, molecular response properties are observable measures of how a molecule changes as a result of interactions with external sources, such as applied electric and/or magnetic fields, nuclear magnetic moments, nuclear displacements from equilibrium, etc. For example, electric dipole polarizability and hyperpolarizability tensors are measures of how a molecule's electric dipole moment changes in response to an external electric field. Molecular response properties of interest are typically origin-independent. However, in many cases a corresponding response property *density* is origin-dependent and a simple definition of an atomic response property as the integral of such a density over the space of the atoms results in origin-dependent atomic contributions. This origin-dependence arises from a ''null'' molecular property that vanishes for the whole molecule but not for an atom in a molecule. The classic example is the electric dipole moment (a first-order response to an external electric field) for neutral molecules, in which case the ''null'' molecular property is the net charge. Because property densities are not unique (any function which integrates to zero for the molecule can be added to a ''basic'' property density), one may try and circumvent this problem by defining a property density that is origin-independent or whose atomic integrals are origin-independent. This is not usually possible, however.

A physically meaningful method for defining origin-independent atomic contributions to electric dipoles, electric polarizabilities, magnetizabilities and other response properties was first introduced by Bader et al. [2–10]. The essence of this method is to express each atomic contribution as the sum of an atomic ''polarization'' contribution and a set of ''transfer'' contributions associated with each group to which the atom is bonded. The expression and interpretation of an atomic polarization contribution, as the polarization of a ''null'' property density

distribution within the nuclear-centered atom, is similar to that for the molecule as a whole. The transfer contribution associated with a bonded group to an atomic response property is given by the weighted translation vector from the atom's nucleus to the corresponding bond critical point, the weight being the bonded group's contribution to the appropriate "null" molecular property. This method not only provides transferable atomic and group response properties consistent with additivity schemes developed from experimental data [3–10], but is also quite intuitive. For atoms in one or more rings, however, ambiguities may arise in the definition of a bonded group. It is, therefore, necessary to restate the transfer contributions to an atomic response property more generally, in terms of atomic bond contributions. One well defined and computationally straightforward method for defining such atomic bond contributions to any "null" properties and for any molecular system, including those with rings and cages, is described in this chapter. In the absence of rings and cages, this more general method reduces to the original method of Bader et al. [2–10]. Using this method, expressions for atomic contributions to electric dipole moments, electric static dipole polarizabilities, infrared vibrational absorption intensities, nuclear virial energies, and magnetizabilities are presented in some detail, with some illustrative numerical examples.

Although inclusion of nuclear virial energies in a list of response properties may seem out of place, a discussion of these is included because many molecular response properties of interest explicitly involve energy derivatives with respect to nuclear positions (e.g. vibrational frequencies) or involve non-stationary point geometries (e.g., vertical transition energies) and, therefore, a well defined atomic contribution to the total molecular energy (i.e. electronic energy plus nuclear virial energy) is necessary to determine such atomic response properties. In addition, a working definition of an atomic contribution to the molecular nuclear virial energy seems to require overcoming origin-dependence issues similar to those encountered with the other response properties described here.

3.2
Apparent Origin-dependence of Some Atomic Response Properties

When the electric dipole moment $\boldsymbol{\mu}$ of a molecule is simply partitioned into atomic contributions [2], the atomic contribution, $\underline{\boldsymbol{\mu}}(\Omega)$, for atom Ω consists of an origin-independent "polarization" term, $\boldsymbol{\mu}_{\mathrm{p}}(\Omega)$, and an origin-dependent "charge transfer" term, $\underline{\boldsymbol{\mu}}_{\mathrm{c}}(\Omega)$:

$$\underline{\boldsymbol{\mu}}(\Omega) = -\int_{\Omega} [\mathbf{r} - \mathbf{R}_{\Omega}]\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + [\mathbf{R}_{\Omega} - \mathbf{R}_0]\mathrm{Q}(\Omega) = \boldsymbol{\mu}_{\mathrm{p}}(\Omega) + \boldsymbol{\mu}_{\mathrm{c}}(\Omega) \tag{1}$$

where $\mathrm{Q}(\Omega)$ is the net charge of atom Ω, $\mathbf{R}_{\Omega}$ is the position vector of atom Ω, and $\mathbf{R}_0$ is the arbitrary origin of the molecular coordinate system. Note that underlines are used here to indicate origin-dependent terms.

Because $\underline{\boldsymbol{\mu}}_c(\Omega)$ is dependent on $\mathbf{R}_0$ (unless $Q(\Omega)$ is zero), $\underline{\boldsymbol{\mu}}(\Omega)$ as defined in Eq. (1) is not generally meaningful, just as $\boldsymbol{\mu}$ itself is not usually meaningful for charged molecules.

Laidig and Bader [3] first showed that for neutral molecules, a more sophisticated definition for the "charge transfer" atomic contribution can be used, $\boldsymbol{\mu}_c(\Omega)$, one which is both origin-independent and physically reasonable:

$$\boldsymbol{\mu}_c(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_b(\Omega|\Lambda) - \mathbf{R}_\Omega] Q({\Lambda_\Omega}^G) \tag{2}$$

In this expression $N_b(\Omega)$ is the number of bond critical points connected to the nucleus of atom Ω, $\mathbf{R}_b(\Omega|\Lambda)$ is the position vector of the bond critical point between atom Ω and atom Λ, and $Q({\Lambda_\Omega}^G)$ is the net charge of the group ${\Lambda_\Omega}^G$ that is bonded to atom Ω via atom Λ. For example, in the methanol molecule $\mathbf{R}_b(C|O)$ is the position vector of the bond critical point between the C and O atoms and $Q({O_C}^G)$ is the net charge of the OH group.

Formulas analogous to Eq. (2) have also been used to partition polarizability tensors [3, 4], magnetizability tensors [5–7], and electronic transition moments [8] of molecules into origin-independent atomic and group contributions.

For molecular structures in which atoms Ω and Λ are bonded within one or more rings, however, $Q({\Lambda_\Omega}^G)$ is usually not well defined (the exception being symmetrically equivalent bonds connecting the atom within the ring or rings) because the group ${\Lambda_\Omega}^G$ to which atom Λ belongs has at least one other bond to atom Ω, say via atom Λ', and it is thus ambiguous which portion of the group containing both Λ and Λ' is defined as ${\Lambda_\Omega}^G$ and which is defined as ${\Lambda'}_\Omega^G$. While for atoms in single rings one may attempt to use the well defined ring critical point instead of the two corresponding bond critical points in Eq. (2), for atoms in more than one ring this will not work and one is again left with an ambiguous definition for ${\Lambda_\Omega}^G$.

Contrary to what has been stated previously [6, 8–10], for isolated molecules this problem *cannot* be solved by using Gauss' law to express $Q({\Lambda_\Omega}^G)$ as the electric field flux through the interatomic surface between atom Λ and atom Ω.

$$\boldsymbol{\mu}_c(\Omega) \neq (-1/4\pi) \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_b(\Omega|\Lambda) - \mathbf{R}_\Omega] \oint\!\!\int d\mathbf{S}(\Lambda|\Omega; \mathbf{r}_s) \bullet \mathbf{E}(\mathbf{r}_s)$$

(for isolated molecules) (3)

The inequality in Eq. (3) ultimately arises because the net charge of an atom in a molecule cannot be solely expressed in terms of electric field fluxes through its interatomic surfaces unless the atom is finite in size, i.e. unless the union of its interatomic surfaces defines a surface completely enclosing a finite volume for the atom:

$$Q(\Omega) = (-1/4\pi) \oint d\mathbf{S}(\Omega|\mathrm{Inf}; \mathbf{r}_s) \bullet \mathbf{E}(\mathbf{r}_s) + (-1/4\pi) \sum_{\Lambda=1}^{\mathrm{Nb}(\Omega)} \oint d\mathbf{S}(\Omega|\Lambda; \mathbf{r}_s) \bullet \mathbf{E}(\mathbf{r}_s) \quad (4)$$

If a portion $S(\Omega|\mathrm{Inf})$ of an atom's bounding surface extends to infinity, the first term on the RHS of Eq. (4) will usually be non-zero. Unlike local property density functions (e.g., the electron density and its derivatives, the stress tensor, the current density, etc.), which decay exponentially far from a molecule, the electric field is a non-local function that decays only as r^{-2} far from a molecule. Because a surface grows in area as r^2, the contribution to the net electric field flux through the infinite portions of an atomic surface will not vanish.

To extend Eq. (2) to systems with rings and cages it seems useful, if not necessary, to dispense with the notion of the bonded group ${\Lambda_\Omega}^G$ and simply restate Eq. (2) more simply, in terms of bond contributions:

$$\boldsymbol{\mu}_c(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] Q(\Omega|\Lambda) \quad (5)$$

where $Q(\Omega|\Lambda)$ is now defined as the contribution from the *directed* bond between atoms Ω and Λ to the net charge of atom Ω. The problem is then to be able to uniquely determine a set of such bond charges $Q(\Omega|\Lambda)$ in any molecule to enable use of Eq. (5). In the absence of a ring containing Ω and Λ, $Q(\Omega|\Lambda)$ should equal $Q({\Lambda_\Omega}^G)$, so Eq. (2) is recovered in those instances.

The situation is similar for atomic polarizability, magnetizability and other response properties.

In the following section, a method of partitioning any "null" (zero value) molecular property η of any molecular structure into unique bond contributions, $\eta(\Omega|\Lambda)$, is described.

3.3
Bond Contributions to "Null" Molecular Properties

Consider any "null" (zero value) property η of a molecule. Examples of η are:

- the net charge, Q, of a neutral molecule
- the change in net charge, δQ, of a molecule due to a perturbation such as an electric field or nuclear displacement
- the net current vector, $\langle \mathbf{J} \rangle$, induced in a molecule by an external magnetic field (when the continuity equation is satisfied)
- the net Ehrenfest force, $\langle \mathbf{F}_E \rangle$, acting on the electrons
- the energy-gradient-based force, $\mathbf{G}_A$, acting on a nucleus A at an equilibrium or other stationary point geometry

- the Hellman–Feynman electrostatic force, $\mathbf{F}_A$, acting on a nucleus A at an equilibrium geometry or other stationary point geometry (when the Hellman–Feynman electrostatic theorem is satisfied)
- the sum of the Hellman–Feynman electrostatic forces, $\Sigma(\mathbf{F}_A)$, acting on the nuclei at any geometry (when the Hellman–Feynman electrostatic theorem is satisfied)
- the sum of the energy-gradient-based forces, $\Sigma(\mathbf{G}_A)$, acting on the nuclei at any geometry

Although such properties vanish for molecules, their atomic contributions $\eta(\Omega)$ and corresponding bond contributions $\eta(\Omega|\Lambda)$ will not typically vanish. The contributions $\eta(\Omega)$ and $\eta(\Omega|\Lambda)$ may be of interest directly and/or they may appear in expressions for other atomic properties that are of interest. Equations (2) and (5) are examples of the latter.

While the atomic property $\eta(\Omega)$ is typically well defined and easily calculated by simple partitioning of η (as in each of the cases listed above), a method for calculating the corresponding bond contributions $\eta(\Omega|\Lambda)$ has not been previously defined in a general and computationally convenient way.

Considering a molecule with N_a atoms, η is a sum of N_a atomic contributions $\eta(\Omega)$:

$$\eta = \sum_{\Omega=1}^{N_a} \eta(\Omega) = 0 \tag{6}$$

Using QTAIM [1] to uniquely identify the bonds, rings and cages in a molecule, each atomic contribution $\eta(\Omega)$ to η can be expressed as a sum of bond contributions $\eta(\Omega|\Lambda)$:

$$\eta(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} \eta(\Omega|\Lambda) \tag{7}$$

where $N_b(\Omega)$ is the number of bond critical points to atom Ω and $\eta(\Omega|\Lambda)$ is the contribution to $\eta(\Omega)$ from the *directed* bond between atoms Ω and atom Λ. To uniquely solve for all of the $2N_b$ bond contributions in a molecule, the following constraint (meaningful only for "null" molecular properties) is used for each bond:

$$\eta(\Omega|\Lambda) + \eta(\Lambda|\Omega) = 0 \tag{8}$$

In the absence of rings, the set of $N_a + N_b$ equations defined by Eqs (7) and (8) is sufficient to uniquely determine the $2N_b = N_a + N_b - 1$ bond contributions because one of the equations is redundant, due to the constraint of Eq. (6). If the

molecule contains one or more rings, however, an additional constraint is necessary for each ring. The following constraint is proposed here:

$$\sum_{\Omega=1}^{N_a(R)} \eta(\Omega|\Omega+1) = 0 \tag{9}$$

where $N_a(R)$ is the number of atoms in ring R and $\eta(\Omega|\Omega+1)$ is the contribution from the bond between the Ωth atom in the ring and the next atom in the ring. Note that the counting is done *one-way* (clockwise or counterclockwise) but that either one-way sum is equivalent, because of Eq. (8). Note, however, that Eq. (9) itself does not follow from Eq. (8).

Equation (9) states that the *one-way* sum of the bond contributions from a ring to a "null" molecular property is necessarily zero, because a ring is a closed circuit of bonds. For a molecule with N_r rings, Eqs (7)–(9) constitute a set of $N_a + N_b + N_r$ linear equations involving $2N_b$ unknowns. There is, however, one redundant equation, because of the constraint of Eq. (6) and also one redundant ring equation for each cage. For a molecule with N_c cages there are, therefore, $N_a + N_b + N_r - 1 - N_c$ linearly independent equations. Using the Poincaré–Hopf relationship between N_a, N_b, N_r, and N_c [1]:

$$\begin{aligned}\text{Number of independent equations} &= N_a + N_b + N_r - 1 - N_c = 2N_b \\ &= \text{number of bond contributions}\end{aligned} \tag{10}$$

Thus, the set of linear equations defined by Eqs (7)–(9) has a unique solution for any structure and any "null" property and can easily be solved without manual effort using a linear equation solver such as the DGELSS routine of LAPACK [11], which enables *apparently* (but only apparently) overdetermined sets of equations as input.

Figures 3.1–3.8 show the bond contributions $\eta(\Omega|\Lambda)$ to a generic "null" molecular property η for some generic molecular structures, in terms of the atomic contributions to η. Although an understanding, or picture, of the bond contributions $\eta(\Omega|\Lambda)$ in terms of atomic contributions is not necessary to use them, occasionally such pictures may be helpful for interpretation.

The result obtained from Eqs (7)–(9) for a bond contribution $\eta(\Omega|\Lambda)$ in a ring system is equivalent to the result of averaging the corresponding bond contribu-

A—B

$\eta(A|B) = \eta(A)$

$\eta(B|A) = \eta(B)$

Fig. 3.1 Bond contributions to η in a diatomic molecule.

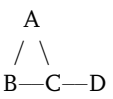
A—B—C

$$\eta(\mathrm{A}|\mathrm{B}) = \eta(\mathrm{A})$$
$$\eta(\mathrm{B}|\mathrm{A}) = -\eta(\mathrm{A})$$
$$\eta(\mathrm{B}|\mathrm{C}) = -\eta(\mathrm{C})$$
$$\eta(\mathrm{C}|\mathrm{B}) = \eta(\mathrm{C})$$

Fig. 3.2 Bond Contributions to η in a triatomic chain molecule.

$$\eta(\mathrm{A}|\mathrm{B}) = -(1/3)[2\eta(\mathrm{B}) + \eta(\mathrm{C})]$$
$$\eta(\mathrm{A}|\mathrm{C}) = -(1/3)[2\eta(\mathrm{C}) + \eta(\mathrm{B})]$$
$$\eta(\mathrm{B}|\mathrm{A}) = -(1/3)[2\eta(\mathrm{A}) + \eta(\mathrm{C})]$$
$$\eta(\mathrm{B}|\mathrm{C}) = -(1/3)[2\eta(\mathrm{C}) + \eta(\mathrm{A})]$$
$$\eta(\mathrm{C}|\mathrm{A}) = -(1/3)[2\eta(\mathrm{A}) + \eta(\mathrm{B})]$$
$$\eta(\mathrm{C}|\mathrm{B}) = -(1/3)[2\eta(\mathrm{B}) + \eta(\mathrm{A})]$$

Fig. 3.3 Bond contributions to η in a triatomic ring molecule.



$$\eta(\mathrm{A}|\mathrm{B}) = -(1/3)[2\eta(\mathrm{B}) + \eta(\mathrm{C}) + \eta(\mathrm{D})]$$
$$\eta(\mathrm{A}|\mathrm{C}) = -(1/3)[2\eta(\mathrm{C}) + 2\eta(\mathrm{D}) + \eta(\mathrm{B})]$$
$$\eta(\mathrm{B}|\mathrm{A}) = -(1/3)[2\eta(\mathrm{A}) + \eta(\mathrm{C}) + \eta(\mathrm{D})]$$
$$\eta(\mathrm{B}|\mathrm{C}) = -(1/3)[2\eta(\mathrm{C}) + 2\eta(\mathrm{D}) + \eta(\mathrm{A})]$$
$$\eta(\mathrm{C}|\mathrm{A}) = -(1/3)[2\eta(\mathrm{A}) + \eta(\mathrm{B})]$$
$$\eta(\mathrm{C}|\mathrm{B}) = -(1/3)[2\eta(\mathrm{B}) + \eta(\mathrm{A})]$$
$$\eta(\mathrm{C}|\mathrm{D}) = -\eta(\mathrm{D})$$
$$\eta(\mathrm{D}|\mathrm{C}) = \eta(\mathrm{D})$$

Fig. 3.4 Bond contributions to η in a branched triatomic ring molecule.

A—B—C—D

$$\eta(\mathrm{A}|\mathrm{B}) = \eta(\mathrm{A})$$

$$\eta(\mathrm{B}|\mathrm{A}) = -\eta(\mathrm{A})$$

$$\eta(\mathrm{B}|\mathrm{C}) = -\eta(\mathrm{C}) - \eta(\mathrm{D})$$

$$\eta(\mathrm{C}|\mathrm{B}) = -\eta(\mathrm{A}) - \eta(\mathrm{B})$$

$$\eta(\mathrm{C}|\mathrm{D}) = -\eta(\mathrm{D})$$

$$\eta(\mathrm{D}|\mathrm{C}) = \eta(\mathrm{D})$$

Fig. 3.5 Bond contributions to η in a tetraatomic chain molecule.

```
    C
   /
A—B
   \
    D
```

$$\eta(\mathrm{A}|\mathrm{B}) = \eta(\mathrm{A})$$

$$\eta(\mathrm{B}|\mathrm{A}) = -\eta(\mathrm{A})$$

$$\eta(\mathrm{B}|\mathrm{C}) = -\eta(\mathrm{C})$$

$$\eta(\mathrm{B}|\mathrm{D}) = -\eta(\mathrm{D})$$

$$\eta(\mathrm{C}|\mathrm{B}) = \eta(\mathrm{C})$$

$$\eta(\mathrm{D}|\mathrm{B}) = \eta(\mathrm{D})$$

Fig. 3.6 Bond contributions to η in a tetraatomic branched chain molecule.

tion resulting from every possible way of (artificially) *minimally* opening the ring (or rings) in which A and B are connected such that they are no longer connected in a ring or rings. Figures 3.9 and 3.10 illustrate this for a triatomic ring and a tetraatomic bicyclic ring.

Although a procedure for determining $\eta(\Omega|\Lambda)$ on the basis of such pictures would obviously be tedious, the pictures do provide some insight into the meaning of the ring constraint defined by Eq. (9).

As shown in Section 3.2 for electric dipole moments, one of the uses of the decomposition of "null" molecular properties into bond contributions is that apparently origin-dependent atomic contributions to *other* molecular properties, such as electric dipole moments, electric polarizabilities, magnetizabilities, infrared vibrational absorption intensities, and nuclear virial energies, can be defined in an origin-independent and physically reasonable manner.

$$\eta(\mathrm{A}|\mathrm{B}) = -(1/4)[3\eta(\mathrm{B}) + 2\eta(\mathrm{C}) + \eta(\mathrm{D})]$$

$$\eta(\mathrm{A}|\mathrm{D}) = -(1/4)[3\eta(\mathrm{D}) + 2\eta(\mathrm{C}) + \eta(\mathrm{B})]$$

$$\eta(\mathrm{B}|\mathrm{A}) = -(1/4)[3\eta(\mathrm{A}) + 2\eta(\mathrm{D}) + \eta(\mathrm{C})]$$

$$\eta(\mathrm{B}|\mathrm{C}) = -(1/4)[3\eta(\mathrm{C}) + 2\eta(\mathrm{D}) + \eta(\mathrm{A})]$$

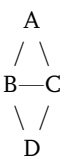
$$\eta(\mathrm{C}|\mathrm{B}) = -(1/4)[3\eta(\mathrm{B}) + 2\eta(\mathrm{A}) + \eta(\mathrm{D})]$$

$$\eta(\mathrm{C}|\mathrm{D}) = -(1/4)[3\eta(\mathrm{D}) + 2\eta(\mathrm{A}) + \eta(\mathrm{B})]$$

$$\eta(\mathrm{D}|\mathrm{A}) = -(1/4)[3\eta(\mathrm{A}) + 2\eta(\mathrm{B}) + \eta(\mathrm{C})]$$

$$\eta(\mathrm{D}|\mathrm{C}) = -(1/4)[3\eta(\mathrm{C}) + 2\eta(\mathrm{B}) + \eta(\mathrm{A})]$$

Fig. 3.7 Bond contributions to η in a tetraatomic ring molecule.



$$\eta(\mathrm{A}|\mathrm{B}) = -(1/8)[5\eta(\mathrm{B}) + 3\eta(\mathrm{C}) + 4\eta(\mathrm{D})]$$

$$\eta(\mathrm{A}|\mathrm{C}) = -(1/8)[5\eta(\mathrm{C}) + 3\eta(\mathrm{B}) + 4\eta(\mathrm{D})]$$

$$\eta(\mathrm{B}|\mathrm{A}) = -(1/8)[5\eta(\mathrm{A}) + 2\eta(\mathrm{C}) + \eta(\mathrm{D})]$$

$$\eta(\mathrm{B}|\mathrm{C}) = -(1/4)[2\eta(\mathrm{C}) + \eta(\mathrm{A}) + \eta(\mathrm{D})]$$

$$\eta(\mathrm{B}|\mathrm{D}) = -(1/4)[5\eta(\mathrm{D}) + 2\eta(\mathrm{C}) + \eta(\mathrm{A})]$$

$$\eta(\mathrm{C}|\mathrm{A}) = -(1/8)[5\eta(\mathrm{A}) + 2\eta(\mathrm{B}) + \eta(\mathrm{D})]$$

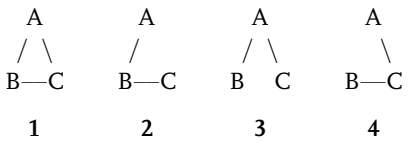
$$\eta(\mathrm{C}|\mathrm{B}) = -(1/4)[2\eta(\mathrm{B}) + \eta(\mathrm{A}) + \eta(\mathrm{D})]$$

$$\eta(\mathrm{C}|\mathrm{D}) = -(1/4)[5\eta(\mathrm{D}) + 2\eta(\mathrm{B}) + \eta(\mathrm{A})]$$

$$\eta(\mathrm{D}|\mathrm{B}) = -(1/8)[5\eta(\mathrm{B}) + 3\eta(\mathrm{C}) + 4\eta(\mathrm{A})]$$

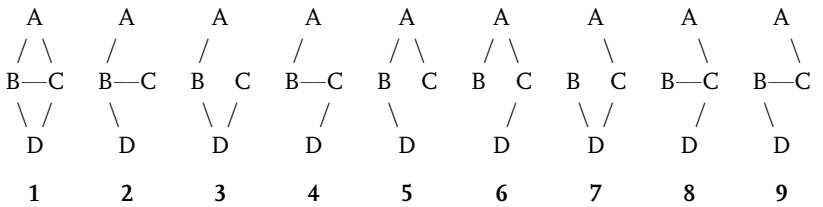
$$\eta(\mathrm{D}|\mathrm{C}) = -(1/8)[5\eta(\mathrm{C}) + 3\eta(\mathrm{B}) + 4\eta(\mathrm{A})]$$

Fig. 3.8 Bond contributions to η in a tetraatomic bicylic molecule.



$$\eta(\mathrm{A}|\mathrm{B})_1 = (1/3)[\eta(\mathrm{A}|\mathrm{B})_2 + \eta(\mathrm{A}|\mathrm{B})_3 + \eta(\mathrm{A}|\mathrm{B})_4]$$
$$= (1/3)[\eta(\mathrm{A}) - \eta(\mathrm{B}) + 0] = -(1/3)[2\eta(\mathrm{B}) + \eta(\mathrm{C})]$$

Fig. 3.9 Bond Contributions to η in a Tetraatomic Bicylic Molecule.



$$\eta(\mathrm{A}|\mathrm{B})_1 = (1/8)[\eta(\mathrm{A}|\mathrm{B})_2 + \eta(\mathrm{A}|\mathrm{B})_3 + \eta(\mathrm{A}|\mathrm{B})_4 + \eta(\mathrm{A}|\mathrm{B})_5 + \eta(\mathrm{A}|\mathrm{B})_6$$
$$+ \eta(\mathrm{A}|\mathrm{B})_7 + \eta(\mathrm{A}|\mathrm{B})_8 + \eta(\mathrm{A}|\mathrm{B})_9]$$
$$= -(1/8)[5\eta(\mathrm{B}) + 3\eta(\mathrm{C}) + 4\eta(\mathrm{D})]$$

Fig. 3.10 Bond Contributions to η in a Tetraatomic Bicylic Molecule.

It should be emphasized that this method is only applicable to ''null'' molecular properties, because of the constraints defined in Eqs (6)–(9). A generalization of Eqs (6)–(9) to the decomposition of any molecular property P (''null'' or ''non-null'') into physically reasonable bond contributions might be possible, and would be highly desirable, but is not considered further here except to note that such a method might produce the following formula for $\mathrm{P}(\Omega|\Lambda)$ when Ω and Λ are not connected in a ring:

$$\mathrm{P}(\Omega|\Lambda) = [\mathrm{P}(\Omega)/(\mathrm{P} - \mathrm{P}(\Omega))]\mathrm{P}(\Lambda_\Omega{}^{\mathrm{G}}) \tag{11}$$

where $\mathrm{P}(\Lambda_\Omega{}^{\mathrm{G}})$ is the net value of P for the group that is bonded to atom Ω via atom Λ.

3.4 Bond Contributions to Atomic Charges in Neutral Molecules

Using Eq. (7), the atomic charge $\mathrm{Q}(\Omega)$ of an atom Ω in a neutral molecule can be expressed in terms of corresponding bond charges $\mathrm{Q}(\Omega|\Lambda)$ as:

Table 3.1 Symmetrically unique atomic charges and their bond contributions in *para*-nitroaniline.

Atom, Ω	Q(Ω)	$Q(\Omega\|\Lambda_1)\Lambda_1$	$Q(\Omega\|\Lambda_2)\Lambda_2$	$Q(\Omega\|\Lambda_3)\Lambda_3$
C1	+0.542	+0.468 N2	+0.037 C3	+0.037 C4
N2	−1.309	−0.468 C1	−0.421 H9	−0.421 H9
C3	+0.047	−0.037 C1	+0.066 C5	+0.018 H7
C5	+0.068	−0.066 C3	−0.063 H12	+0.197 C11
H7	−0.018	−0.018 C3		
H10	+0.421	+0.421 N2		
C11	+0.313	−0.197 C5	−0.197 C6	+0.706 N14
H12	+0.063	+0.063 C5		
N14	+0.376	−0.706 C11	+0.541 O15	+0.541 O16
O15	−0.541	−0.541 N14		
Total[a]	+0.002			

[a] The deviation from a total net charge of zero is because of numerical integration errors for the atoms. The maximum atomic integration error, as measured by |L(Ω)|, was 0.001 a.u. for N14.

$$Q(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} Q(\Omega|\Lambda) \tag{12}$$

where the set of $Q(\Omega|\Lambda)$ are determined by solving the set of linear equations defined by Eqs (7)–(9). As an example, Table 3.1 shows the set of $Q(\Omega)$ and $Q(\Omega|\Lambda)$ for each symmetrically unique atom in *para*-nitroaniline (shown in Fig. 3.11), calculated at HF/6-311++G(2d,2p)//HF/6-311++G(2d,2p). In addition to being useful for partitioning electric dipole moments and their derivatives into atomic contributions, such bond charges might find application in their own right, as descriptors in QSAR studies or in the construction of molecular mechanics force fields, for example.

3.5 Atomic Contributions to Electric Dipole Moments of Neutral Molecules

The electric dipole moment, μ, of a molecule is given by:

$$\boldsymbol{\mu} = -\int [\mathbf{r} - \mathbf{R}_0]\rho(\mathbf{r})\, d\mathbf{r} + \sum_{\Omega=1}^{N_a} Z_\Omega [\mathbf{R}_\Omega - \mathbf{R}_0] \tag{13}$$

where $\rho(\mathbf{r})$ is the electron density at a point $\mathbf{r}$ in real space, Z_Ω is the nuclear charge of atom Ω, $\mathbf{R}_\Omega$ is the position vector of nucleus Ω, and $\mathbf{R}_0$ is the arbitrary

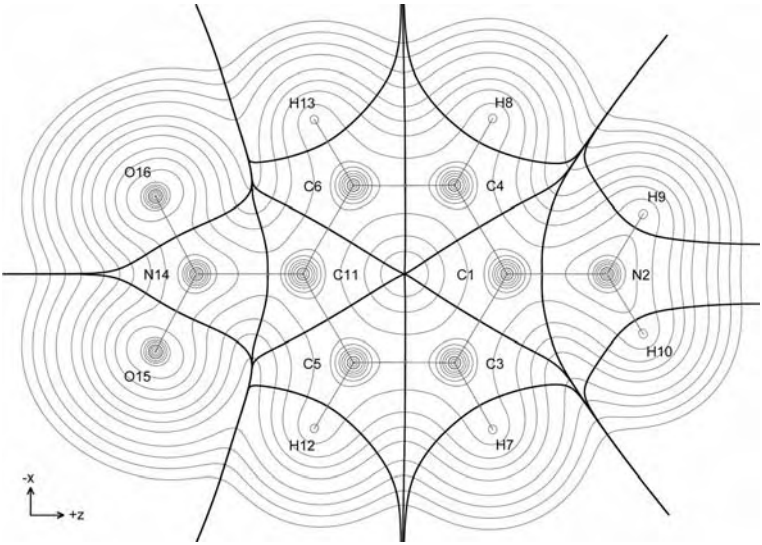


Fig. 3.11 Electron density contours of *para*-nitroaniline in the nuclear plane, overlaid with interatomic surfaces (bold) and bond paths (semi-bold).

origin of the molecular coordinate system. For neutral molecules ($Q = 0$), $\boldsymbol{\mu}$ is independent of $\mathbf{R}_0$.

As shown in Eq. (1), $\boldsymbol{\mu}$ can be expressed [2, 3] in terms of atomic dipole polarization contributions, $\boldsymbol{\mu}_p(\Omega)$, and origin-dependent atomic charge-transfer dipole contributions, $\underline{\boldsymbol{\mu}}_c(\Omega)$:

$$\begin{aligned}
\boldsymbol{\mu} &= \sum_{\Omega=1}^{N_a} \left\{ -\int_\Omega [\mathbf{r} - \mathbf{R}_0]\rho(\mathbf{r})\,d\mathbf{r} + Z_\Omega[\mathbf{R}_\Omega - \mathbf{R}_0] \right\} \\
&= \sum_{\Omega=1}^{N_a} \left\{ -\int_\Omega [\mathbf{r} - \mathbf{R}_\Omega\rho(\mathbf{r})\,d\mathbf{r} + [\mathbf{R}_\Omega - \mathbf{R}_0] \left[\int_\Omega -\rho(\mathbf{r})\,d\mathbf{r} + Z_\Omega \right] \right\} \\
&= \sum_{\Omega=1}^{N_a} \left\{ -\int_\Omega [\mathbf{r} - \mathbf{R}_\Omega]\rho(\mathbf{r})\,d\mathbf{r} + [\mathbf{R}_\Omega - \mathbf{R}_0]Q(\Omega) \right\} \\
&= \sum_{\Omega=1}^{N_a} \{\boldsymbol{\mu}_p(\Omega) + \underline{\boldsymbol{\mu}}_c(\Omega)\} = \sum_{\Omega=1}^{N_a} \underline{\boldsymbol{\mu}}(\Omega) = \boldsymbol{\mu}_p + \boldsymbol{\mu}_c
\end{aligned} \tag{14}$$

where $Q(\Omega)$ is the net charge of atom Ω. Using Eq. (12), the origin-dependent term $\underline{\boldsymbol{\mu}}_c(\Omega)$ can be expressed in terms of bond charges $Q(\Omega|\Lambda)$ as:

$$\underline{\boldsymbol{\mu}}_c(\Omega) = [\mathbf{R}_\Omega - \mathbf{R}_0]Q(\Omega) = [\mathbf{R}_\Omega - \mathbf{R}_0] \sum_{\Lambda=1}^{N_b(\Omega)} Q(\Omega|\Lambda) \tag{15}$$

Using Eq. (8), a corresponding origin-independent expression, $\boldsymbol{\mu}_c(\Omega)$, can be defined, as shown in Eq. (5) and below:

$$\boldsymbol{\mu}_c(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)]Q(\Omega|\Lambda) \tag{16}$$

where $\mathbf{R}_b(\Omega|\Lambda)$ is the position vector of the bond critical point between atoms Ω and Λ.

Note that for each product $\mathbf{R}_b(\Omega|\Lambda)^* Q(\Omega|\Lambda)$ for atom Ω, there is a corresponding term for atom Λ which cancels it, because of Eq. (8), and therefore:

$$\begin{aligned}\sum_{\Omega=1}^{N_a} \boldsymbol{\mu}_c(\Omega) &= \sum_{\Omega=1}^{N_a} \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)]Q(\Omega|\Lambda) \\ &= \sum_{\Omega=1}^{N_a} \mathbf{R}_\Omega Q(\Omega) = \sum_{\Omega=1}^{N_a} \underline{\boldsymbol{\mu}}_c(\Omega)\end{aligned} \tag{17}$$

$$\boldsymbol{\mu} = \sum_{\Omega=1}^{N_a} \boldsymbol{\mu}(\Omega) = \sum_{\Omega=1}^{N_a} \{\boldsymbol{\mu}_p(\Omega) + \boldsymbol{\mu}_c(\Omega)\} = \boldsymbol{\mu}_p + \boldsymbol{\mu}_c \tag{18}$$

Note that all the quantities in the expressions for $\boldsymbol{\mu}_p(\Omega)$ and $\boldsymbol{\mu}_c(\Omega)$ are uniquely determined by the molecular charge distribution. An essential criterion for any atomic property is that, once formally defined, its evaluation should be uniquely determined entirely by the molecular wavefunction, just as the evaluation of the corresponding molecular property is.

As an example, Table 3.2 shows the z-axis component of $\boldsymbol{\mu}_p(\Omega)$ and $\boldsymbol{\mu}_c(\Omega)$ and $\boldsymbol{\mu}(\Omega)$ for each symmetrically unique atom in *para*-nitroaniline (shown in Fig. 3.11), calculated at HF/6-311++G(2d,2p)//HF/6-311++G(2d,2p). The total charge-transfer contribution, $\boldsymbol{\mu}_c$, is over six times that of the opposing total polarization contribution, $\boldsymbol{\mu}_p$. The dominant contributor to $\boldsymbol{\mu}_c$ is the NO_2 nitrogen, N11, especially the contribution from its bond to the ring, $\boldsymbol{\mu}_c(\text{N14|C11})_z$.

3.6 Atomic Contributions to Electric Polarizabilities

When a molecule is placed in an electric field, its electron distribution changes in response (nuclear geometry changes are assumed to be negligible here). A useful measure of this response is the molecular dipole moment $\boldsymbol{\mu}$, whose first derivative and higher derivatives with respect to the electric field correspond to measurable molecular polarizability and hyperpolarizability tensors [13]. Using QTAIM [1], more detailed information about these molecular responses can be obtained, by partitioning the dipole moment and its electric field derivatives into atomic contributions [3, 4].

Table 3.2 Atomic and bond contributions (in a.u.) to the z-axis dipole moment of *para*-nitroaniline.

Atom, Ω	$\mu(\Omega)_z$	$\mu_p(\Omega)_z$	$\mu_c(\Omega)_z$	$\mu_c(\Omega\|\Lambda_1)_z\Lambda_1$	$\mu_c(\Omega\|\Lambda_1)_z\Lambda_2$	$\mu_c(\Omega\|\Lambda_1)_z\Lambda_3$
C1	+0.381	+0.751	−0.370	−0.422 N2	+0.026 C3	+0.026 C4
N2	−0.010	+0.170	−0.180	−0.782 C1	+0.301 H9	+0.301 H10
C3	+0.101	+0.005	+0.096	+0.024 C1	+0.083 C5	−0.011 C6
C5	+0.144	−0.030	+0.174	+0.087 C3	+0.128 C11	−0.042 H12
H7	−0.072	−0.066	−0.006	−0.006 C3		
H10	+0.016	−0.087	+0.103	+0.013 N2		
C11	+0.324	−0.568	+0.892	+0.126 C5	+0.126 C7	+0.641 N14
H12	+0.012	+0.035	−0.023	−0.023 C5		
N14	+0.927	−0.922	+1.849	+1.290 C11	+0.279 O15	+0.279 O16
O15	+0.436	+0.146	+0.290	+0.290 N14		
Total[a,b]	+2.893	−0.563	+3.456			

[a] The analytically calculated value for $\mu(\Omega)_z$ is +2.901 a.u.
[b] An experimental dipole moment for *para*-nitroaniline (measured in acetone) is 2.44 a.u. [12].

The static dipole polarizability tensor of a molecule, $\boldsymbol{\alpha}$, is the gradient ($\nabla_{\mathbf{E}}$) of the molecular electric dipole moment $\boldsymbol{\mu}$ with respect an external, uniform, and time-independent electric field $\mathbf{E}$, evaluated in the limit of zero field strength:

$$\boldsymbol{\alpha} = [\nabla_{\mathbf{E}}\boldsymbol{\mu}]_{\mathrm{E}=0} = [\mathbf{i}(\partial\boldsymbol{\mu}/\partial \mathrm{E_x}) + \mathbf{j}(\partial\boldsymbol{\mu}/\partial \mathrm{E_y}) + \mathbf{k}(\partial\boldsymbol{\mu}/\partial \mathrm{E_z})]_{\mathbf{E}=0} \tag{19}$$

From Eqs (13) and (19), $\boldsymbol{\alpha}$ is given by:

$$\boldsymbol{\alpha} = -\int (\mathbf{r} - \mathbf{R}_0)\rho^{\mathbf{E}}(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{20}$$

where the electron density derivative $\rho^{\mathbf{E}}(\mathbf{r})$ is given by:

$$\rho^{\mathbf{E}}(\mathbf{r}) = [\nabla_{\mathbf{E}}\rho(\mathbf{r})]_{\mathrm{E}=0} = [\mathbf{i}(\partial\rho(\mathbf{r})/\partial \mathrm{E_x}) + \mathbf{j}(\partial\rho(\mathbf{r})/\partial \mathrm{E_y}) + \mathbf{k}(\partial\rho(\mathbf{r})/\partial \mathrm{E_z})]_{\mathrm{E}=0} \tag{21}$$

Throughout this section, the notation $\mathrm{t}^{\mathbf{E}}$ is used to indicate the gradient of the term t with respect to $\mathbf{E}$, evaluated at $\mathbf{E} = 0$. In terms of atomic contributions we have, from Section 3.5:

$$\begin{aligned}\boldsymbol{\alpha} &= \sum_{\Omega=1}^{\mathrm{N_a}} \boldsymbol{\alpha}(\Omega) = \sum_{\Omega=1}^{\mathrm{N_a}} \boldsymbol{\mu}^{\mathbf{E}}(\Omega) = \sum_{\Omega=1}^{\mathrm{N_a}} \{\boldsymbol{\mu}_{\mathrm{p}}{}^{\mathbf{E}}(\Omega) + \boldsymbol{\mu}_{\mathrm{c}}{}^{\mathbf{E}}(\Omega)\} \\ &= \sum_{\Omega=1}^{\mathrm{N_a}} \{\boldsymbol{\alpha}_{\mathrm{p}}(\Omega) + \boldsymbol{\alpha}_{\mathrm{c}}(\Omega)\} = \boldsymbol{\alpha}_{\mathrm{p}} + \boldsymbol{\alpha}_{\mathrm{c}}\end{aligned} \tag{22}$$

The atomic dipole polarization gradient with respect to $\mathbf{E}$, $\boldsymbol{\alpha}_{\mathrm{p}}(\Omega) = \boldsymbol{\mu}_{\mathrm{p}}{}^{\mathbf{E}}(\Omega)$, contains a basin (B) contribution, $\boldsymbol{\alpha}_{\mathrm{p,B}}(\Omega)$, arising from the density gradient $\rho^{\mathbf{E}}(\mathbf{r})$ within the unperturbed atomic basin. It also contains a surface (S) contribution, $\boldsymbol{\alpha}_{\mathrm{p,S}}(\Omega)$, arising from the gradient of the atomic surface S with respect to $\mathbf{E}$:

$$\boldsymbol{\alpha}_{\mathrm{p}}(\Omega) = -\int_{\Omega} (\mathbf{r} - \mathbf{R}_{\Omega})\rho^{\mathbf{E}}(\mathbf{r})\,\mathrm{d}\mathbf{r} + \boldsymbol{\alpha}_{\mathrm{p,S}}(\Omega) = \boldsymbol{\alpha}_{\mathrm{p,B}}(\Omega) + \boldsymbol{\alpha}_{\mathrm{p,S}}(\Omega) \tag{23}$$

Similarly, the gradient of the atomic charge transfer dipole contribution with respect to $\mathbf{E}$, $\boldsymbol{\alpha}_{\mathrm{c}}(\Omega) = \boldsymbol{\mu}_{\mathrm{c}}{}^{\mathbf{E}}(\Omega)$, also contains both basin and surface contributions:

$$\boldsymbol{\alpha}_{\mathrm{c}}(\Omega) = \boldsymbol{\alpha}_{\mathrm{c,B}}(\Omega) + \boldsymbol{\alpha}_{\mathrm{c,S}}(\Omega) \tag{24}$$

$$\boldsymbol{\alpha}_{\mathrm{c,B}}(\Omega) = \sum_{\Lambda=1}^{\mathrm{N_b}(\Omega)} \{[\mathbf{R}_{\Omega} - \mathbf{R}_{\mathrm{b}}(\Omega|\Lambda)]\mathrm{Q_B}{}^{\mathbf{E}}(\Omega|\Lambda) \tag{25}$$

$$\boldsymbol{\alpha}_{\mathrm{c,S}}(\Omega) = \sum_{\Lambda=1}^{\mathrm{N_b}(\Omega)} [\mathbf{R}_{\Omega} - \mathbf{R}_{\mathrm{b}}(\Omega|\Lambda)]\mathrm{Q_S}{}^{\mathbf{E}}(\Omega|\Lambda) - \mathrm{Q}(\Omega|\Lambda)\mathbf{R}_{\mathrm{b}}{}^{\mathbf{E}}(\Omega|\Lambda) \tag{26}$$

Using Eqs (7)–(9), the electric field derivatives of the bond charges, $\mathrm{Q}^{\mathbf{E}}(\Omega|\Lambda)$, are determined from the electric field derivatives of the atomic charges $\mathrm{Q}^{\mathbf{E}}(\Omega)$:

$$\mathrm{Q}^{\mathbf{E}}(\Omega) = -\int_{\Omega} \rho^{\mathbf{E}}(\mathbf{r})\,\mathrm{d}\mathbf{r} + \mathrm{Q_S}{}^{\mathbf{E}}(\Omega) = \mathrm{Q_B}{}^{\mathbf{E}}(\Omega) + \mathrm{Q_S}{}^{\mathbf{E}}(\Omega) \tag{27}$$

The atomic basin and surface bond charge derivative contributions, $\mathrm{Q_B}{}^{\mathbf{E}}(\Omega|\Lambda)$ and $\mathrm{Q_S}{}^{\mathbf{E}}(\Omega|\Lambda)$, can be obtained separately from the corresponding atomic charge derivatives $\mathrm{Q_B}{}^{\mathbf{E}}(\Omega)$ and $\mathrm{Q_S}{}^{\mathbf{E}}(\Omega)$.

The contributions to $\boldsymbol{\alpha}(\Omega)$ from the surface derivatives with respect to $\mathbf{E}$ arise because both the spatial definition of an atom and an atomic dipole moment contribution are determined entirely by the molecular charge distribution, which is of course dependent on $\mathbf{E}$. This is physically sound (it is physically essential), but it also makes the term-by-term evaluation of $\boldsymbol{\alpha}_{\mathrm{p}}(\Omega)$ and $\boldsymbol{\alpha}_{\mathrm{c}}(\Omega)$ more difficult than usual. However, evaluation of $\boldsymbol{\alpha}_{\mathrm{p}}(\Omega)$ and $\boldsymbol{\alpha}_{\mathrm{c}}(\Omega)$ by numerical differentiation (using finite field wavefunctions) is straightforward. For example:

$$\begin{aligned}\boldsymbol{\alpha}_{\mathrm{p}}(\Omega) \bullet \mathbf{k} = {} & 1/2\{[\boldsymbol{\mu}_{\mathrm{p}}(\Omega)]_{\mathbf{E}=\varepsilon\mathbf{k}} - [\boldsymbol{\mu}_{\mathrm{p}}(\Omega)]_{\mathbf{E}=0}\}\varepsilon^{-1} \\ & - 1/2\{[\boldsymbol{\mu}_{\mathrm{p}}(\Omega)]_{\mathbf{E}=-\varepsilon\mathbf{k}} - [\boldsymbol{\mu}_{\mathrm{p}}(\Omega)]_{\mathbf{E}=0}\}\varepsilon^{-1}\end{aligned} \tag{28}$$

Table 3.3 Atomic contributions (in a.u.) to the principal components of the electric polarizability tensor of *para*-nitroaniline.

Atom, Ω	$\alpha(\Omega)_{zz}$	$\alpha_B(\Omega)_{zz}$	$\alpha_p(\Omega)_{zz}$	$\alpha_{p,B}(\Omega)_{zz}$	$\alpha_c(\Omega)_{zz}$	$\alpha_{c,B}(\Omega)_{zz}$
C1	+9.332	+10.623	−3.152	−0.417	+12.484	+11.040
N2	+17.854	+16.963	−0.034	+5.210	+17.888	+11.753
C3	+12.865	+13.133	−2.792	+1.458	+15.657	+11.674
C5	+12.077	+12.220	−4.530	+0.013	+16.607	+12.206
H7	+1.885	+1.670	+0.859	+1.164	+1.026	+0.505
H10	+1.380	+1.364	+0.848	+0.944	+0.532	+0.370
C11	+12.136	+13.844	−0.580	+1.514	+12.716	+12.330
H12	+1.250	+1.004	+0.433	+0.632	+0.817	+0.372
N14	+16.944	+15.656	−3.830	−0.841	+20.774	+16.497
O15	+6.514	+6.310	+4.439	+4.406	+2.074	+1.904
Total[a,b]	+128.214	+128.487	−9.083	+22.803	+137.297	+105.685

Atom, Ω	$\alpha(\Omega)_{xx}$	$\alpha_B(\Omega)_{xx}$	$\alpha_p(\Omega)_{xx}$	$\alpha_{p,B}(\Omega)_{xx}$	$\alpha_c(\Omega)_{xx}$	$\alpha_{c,B}(\Omega)_{xx}$
C1	+9.756	+9.830	−4.362	+0.041	+14.118	+9.789
N2	+4.413	+4.435	−1.590	+1.799	+6.002	+2.637
C3	+8.125	+8.385	−3.283	+1.424	+11.408	+6.961
C5	+7.263	+7.768	−3.220	+1.178	+10.483	+6.591
H7	+3.998	+3.558	+0.881	+1.972	+3.117	+1.587
H10	+1.258	+1.221	+0.645	+0.778	+0.613	+0.442
C11	+9.994	+9.663	−3.354	+0.637	+13.348	+9.026
H12	+3.260	+2.936	+0.845	+1.673	+2.415	+1.263
N14	+6.975	+7.367	−1.430	−0.494	+8.404	+7.861
O15	+7.672	+7.578	+2.925	+3.367	+4.747	+4.211
Total[a,b]	+94.291	+94.187	−13.150	+22.766	+107.441	+71.421

where the average over numerical derivatives for both positive and negative fields is used. Numerical differentiation to obtain $\boldsymbol{\alpha}_p(\Omega)$ and $\boldsymbol{\alpha}_c(\Omega)$, together with the straightforward evaluation of the basin contributions (which involve the analytical density derivatives $\rho^{\mathbf{E}}$) enables the surface derivative contributions to $\boldsymbol{\alpha}_p(\Omega)$ and $\boldsymbol{\alpha}_c(\Omega)$ to be obtained by difference. Note that the sum of all of the atomic surface derivative contributions is zero:

$$\boldsymbol{\alpha}_S = \sum_{\Omega=1}^{N_a} \boldsymbol{\alpha}_S(\Omega) = \sum_{\Omega=1}^{N_a} [\boldsymbol{\alpha}_{p,S}(\Omega) + \boldsymbol{\alpha}_{c,S}(\Omega)] = 0 \tag{29}$$

Table 3.3 (*continued*)

Atom, Ω	$\alpha(\Omega)_{yy}$	$\alpha_B(\Omega)_{yy}$	$\alpha_p(\Omega)_{yy}$	$\alpha_{p,B}(\Omega)_{yy}$	$\alpha_c(\Omega)_{yy}$	$\alpha_{c,B}(\Omega)_{yy}$
C1	+3.183	+3.291	+3.185	+3.291	−0.002	0.000
N2	+6.950	+6.936	+6.983	+6.936	−0.033	0.000
C3	+6.337	+6.282	+6.340	+6.282	−0.003	0.000
C5	+4.710	+4.796	+4.742	+4.796	−0.032	0.000
H7	+1.115	+1.113	+1.117	+1.113	−0.002	0.000
H10	+0.389	+0.403	+0.375	+0.403	+0.013	0.000
C11	+5.136	+4.853	+5.113	+4.853	+0.023	0.000
H12	+0.900	+0.904	+0.889	+0.904	+0.011	0.000
N14	+2.190	+2.203	+2.140	+2.203	+0.050	0.000
O15	+3.497	+3.586	+3.504	+3.586	−0.007	0.000
Total[a,b]	+51.354	+51.451	+51.354	+51.451	0.000	

[a] The analytically calculated values of α_{zz}, α_{xx}, and α_{yy} are +128.487, +94.187, and +51.451 a.u.
[b] The experimentally measured isotropic polarizability (measured in acetone) is +114.73 a.u. [12] compared with the analytically calculated isotropic value of +91.375 a.u.

Thus, the molecular polarizability $\boldsymbol{\alpha}$ can be written as:

$$\boldsymbol{\alpha} = \sum_{\Omega=1}^{N_a} \boldsymbol{\alpha}(\Omega) = \sum_{\Omega=1}^{N_a} [\boldsymbol{\alpha}_p(\Omega) + \boldsymbol{\alpha}_c(\Omega)]$$

$$= \sum_{\Omega=1}^{N_a} [\boldsymbol{\alpha}_{p,B}(\Omega) + \boldsymbol{\alpha}_{c,B}(\Omega) + \boldsymbol{\alpha}_{p,S}(\Omega) + \boldsymbol{\alpha}_{c,S}(\Omega)] = \boldsymbol{\alpha}_B + \boldsymbol{\alpha}_S = \boldsymbol{\alpha}_B \tag{30}$$

That the sum of the atomic surface derivative contributions vanishes does not necessarily mean they should be ignored, however, especially when comparing and/or transferring atomic polarizabilities between molecules.

Using the partitioning method outlined above, Bader et al. [4] studied, among other things, the atomic and group contributions to the polarizability tensors of the normal alkane series and showed that the empirical additivity relationship for the isotropic polarizability of normal alkanes in terms of transferable methylene and methyl group contributions is mirrored by the corresponding theoretical group contributions, thus providing a theoretical basis for the empirical relationship or an experimental validation for the theoretical definition of QTAIM polarizabilities.

As an example here, Table 3.3 shows the diagonal components of $\boldsymbol{\alpha}_p(\Omega)$ and $\boldsymbol{\alpha}_c(\Omega)$ and $\boldsymbol{\alpha}(\Omega)$ for each symmetrically unique atom of *para*-nitroaniline, calculated using the coupled–perturbed HF/6-311++G(2d,2p)//HF/6-311++G(2d,2p) level of theory. Also shown in Table 3.3 are the corresponding basin contributions

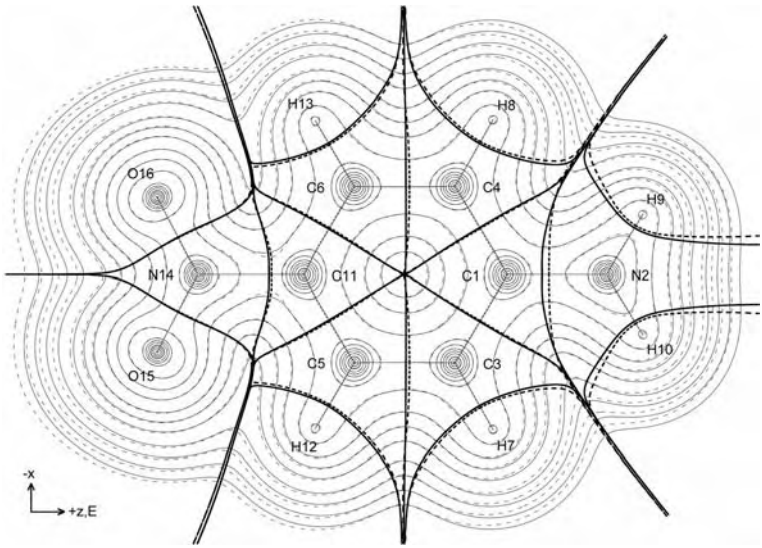


Fig. 3.12 Effect of a finite (0.05 a.u.) electric field applied along the z-axis of *para*-nitroaniline in terms of electron density contours, interatomic surfaces (bold), and bond paths (semi-bold) in the nuclear plane. Dotted lines correspond to the presence of the field whereas solid lines correspond to the absence of the field.

$\boldsymbol{\alpha}_{\mathrm{p,B}}(\Omega)$, $\boldsymbol{\alpha}_{\mathrm{c,B}}(\Omega)$, and $\boldsymbol{\alpha}_{\mathrm{B}}(\Omega)$. The magnitude of the total charge-transfer contribution to $\boldsymbol{\alpha}_{zz}$ is over fourteen times that of the opposing total polarization contribution. The polarizability $\boldsymbol{\alpha}_{xx}$ is similarly dominated by charge transfer whereas the much smaller $\boldsymbol{\alpha}_{yy}$ is due entirely to atomic polarization, the charge-transfer contribution vanishing by symmetry (the non-zero values shown for $\boldsymbol{\alpha}_{\mathrm{c}}(\Omega)_{yy}$ provide a measure of the error in the results for the other atomic polarizability contributions). That the sum of $\boldsymbol{\alpha}(\Omega)_{zz}$ does not equal the sum of $\boldsymbol{\alpha}_{\mathrm{B}}(\Omega)_{zz} = \boldsymbol{\alpha}_{zz}$ is because of error in the numerical differentiation used to obtain the total (basin + surface) $\boldsymbol{\alpha}(\Omega)_{zz}$ contributions, whereas the $\boldsymbol{\alpha}_{\mathrm{B}}(\Omega)_{zz}$ contributions were calculated from the analytical density derivative. Similar statements apply for the yy and zz components. Figure 3.12 shows the effect on the charge distribution of a finite external electric field (0.05 a.u.) applied along the z-axis. Note that the shift in electronic charge is opposed to the electric field whereas the shift in the interatomic surfaces is in the opposite direction [3].

3.7
Atomic Contributions to Vibrational Infrared Absorption Intensities

Within the (double) harmonic approximation, the intensity of absorption, I, of a peak in an infrared spectrum for a molecule is proportional to the first derivative

of the electric dipole moment, $\boldsymbol{\mu}$, of the molecule with respect to one of the normal-mode vibrational coordinates ξ (a combination of nuclear displacement coordinates from an equilibrium geometry), in the limit of $\xi = 0$. For a normal mode of vibration with frequency ν_ξ and normal coordinate ξ, the corresponding theoretical absorption intensity, I_ξ, is [13, 14]:

$$I_\xi = C|(d\boldsymbol{\mu}/d\xi)_{\xi=0}|^2 = C|\boldsymbol{\mu}^\xi|^2 \tag{31}$$

where C is a collection of constants and the notation t^ξ is used to indicate the derivative of the term t with respect ξ, evaluated at $\xi = 0$. Absorption intensities of vibrational spectra can be understood in terms of atomic and bond contributions using the equations in Section 3.5 with Eq. (31). The formalism is similar to that for polarizabilities:

$$\boldsymbol{\mu}^\xi = \sum_{\Omega=1}^{N_a} \boldsymbol{\mu}^\xi(\Omega) = \sum_{\Omega=1}^{N_a} \{\boldsymbol{\mu}_p{}^\xi(\Omega) + \boldsymbol{\mu}_c{}^\xi(\Omega)\} = \boldsymbol{\mu}_p{}^\xi + \boldsymbol{\mu}_c{}^\xi \tag{32}$$

$$\begin{aligned}\boldsymbol{\mu}_p{}^\xi(\Omega) &= \left[-\int_\Omega (\mathbf{r} - \mathbf{R}_\Omega)\rho^\xi(\mathbf{r})\,d\mathbf{r} + \mathbf{R}_\Omega{}^\xi Q(\Omega)\right] + \boldsymbol{\mu}_{p,S}{}^\xi(\Omega) \\ &= \boldsymbol{\mu}_{p,B}{}^\xi(\Omega) + \boldsymbol{\mu}_{p,S}{}^\xi(\Omega)\end{aligned} \tag{33}$$

$$\begin{aligned}\boldsymbol{\mu}_c{}^\xi(\Omega) &= \sum_{\Lambda=1}^{N_b(\Omega)} \{[\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)]Q^\xi(\Omega|\Lambda) + \mathbf{R}_\Omega{}^\xi Q(\Omega|\Lambda) - Q(\Omega|\Lambda)\mathbf{R}_b{}^\xi(\Omega|\Lambda)\} \\ &= \sum_{\Lambda=1}^{N_b(\Omega)} \{[\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)]Q_B{}^\xi(\Omega|\Lambda) + \mathbf{R}_\Omega{}^\xi Q(\Omega|\Lambda)\} \\ &\quad + \sum_{\Lambda=1}^{N_b(\Omega)} \{[\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)]Q_S{}^\xi(\Omega|\Lambda) - Q(\Omega|\Lambda)\mathbf{R}_b{}^\xi(\Omega|\Lambda)\} \\ &= \boldsymbol{\mu}_{c,B}{}^\xi(\Omega) + \boldsymbol{\mu}_{c,S}{}^\xi(\Omega)\end{aligned} \tag{34}$$

$$\boldsymbol{\mu}_S{}^\xi = \sum_{\Omega=1}^{N_a} \boldsymbol{\mu}_S{}^\xi(\Omega) = \sum_{\Omega=1}^{N_a} [\boldsymbol{\mu}_{p,S}{}^\xi(\Omega) + \boldsymbol{\mu}_{c,S}{}^\xi(\Omega)] = 0 \tag{35}$$

$$\begin{aligned}\boldsymbol{\mu}^\xi &= \sum_{\Omega=1}^{N_a} \boldsymbol{\mu}^\xi(\Omega) = \sum_{\Omega=1}^{N_a} [\boldsymbol{\mu}_p{}^\xi(\Omega) + \boldsymbol{\mu}_c{}^\xi(\Omega)] \\ &= \sum_{\Omega=1}^{N_a} [\boldsymbol{\mu}_{p,B}{}^\xi(\Omega) + \boldsymbol{\mu}_{c,B}{}^\xi(\Omega) + \boldsymbol{\mu}_{p,S}{}^\xi(\Omega) + \boldsymbol{\mu}_{c,S}{}^\xi(\Omega)] \\ &= \boldsymbol{\mu}_B{}^\xi + \boldsymbol{\mu}_S{}^\xi = \boldsymbol{\mu}_B{}^\xi\end{aligned} \tag{36}$$

The normal coordinate derivatives of the bond charges $Q^{\xi}(\Omega|\Lambda)$ and their basin and surface contributions, $Q_B{}^{\xi}(\Omega|\Lambda)$ and $Q_S{}^{\xi}(\Omega|\Lambda)$, are determined from the corresponding atomic charge derivatives $Q^{\xi}(\Omega)$ using Eqs (7)–(9). $Q^{\xi}(\Omega)$ is given by:

$$Q^{\xi}(\Omega) = -\int_{\Omega} \rho^{\xi}(\mathbf{r})\, d\mathbf{r} + Q_S{}^{\xi}(\Omega) = Q_B{}^{\xi}(\Omega) + Q_S{}^{\xi}(\Omega) \tag{37}$$

When compared with the corresponding polarizability term $\boldsymbol{\alpha}_p(\Omega)$, the term $\boldsymbol{\mu}_p{}^{\xi}(\Omega)$ contains an additional contribution $-\mathbf{R}_{\Omega}{}^{\xi}Q(\Omega)$, which is the atomic charge-weighted derivative of the nuclear position $\mathbf{R}_{\Omega}$ with respect to ξ_k. Similarly, the charge-transfer term $\boldsymbol{\mu}_c{}^{\xi}(\Omega)$ contains an extra contribution for each bond, $\mathbf{R}_{\Omega}{}^{\xi}Q(\Omega|\Lambda)$, which is the bond-charge-weighted derivative of the nuclear position $\mathbf{R}_{\Omega}$ with respect to coordinate ξ.

Like the atomic polarizability contributions, the term-by-term evaluation of the atomic polarization and charge-transfer contributions to an electric dipole derivative with respect to a normal vibrational coordinate is complicated (but in a physically necessary way) by the dependence of the atomic surfaces on the nuclear positions. As with the polarizabilities, however, the total $\boldsymbol{\mu}_p{}^{\xi}(\Omega)$ and $\boldsymbol{\mu}_c{}^{\xi}(\Omega)$ can each be easily evaluated by numerical differentiation and the terms involving the analytical density derivative $\rho^{\xi}(\mathbf{r})$ can be evaluated directly, thus enabling the surface derivative contributions to be obtained by difference.

Bader et al. [2] were the first to study molecular dipole moment changes associated with normal mode vibrations using QTAIM. Their analysis used finite displacements along a normal coordinate or symmetry coordinate to observe changes in $\boldsymbol{\mu}_p(\Omega)$ and an origin-dependent charge-transfer contribution which is somewhat different, and less generally applicable, than the $\boldsymbol{\mu}_c(\Omega)$ given here, but their results clearly showed that changes in atomic polarizations, changes in atomic charges, and displacement of atomic charges are all important in understanding dipole moment derivatives with respect to normal mode vibrational coordinates, and hence IR spectra.

As an example here, Table 3.4 shows the atomic contributions $\boldsymbol{\mu}^{\xi}(\Omega)$, $\boldsymbol{\mu}_p{}^{\xi}(\Omega)$, and $\boldsymbol{\mu}_c{}^{\xi}(\Omega)$ to the molecular dipole derivative $\boldsymbol{\mu}^{\xi}$ for one of the pair of degenerate bending normal modes of vibration in CO_2 and the asymmetric stretching normal mode, calculated at the coupled–perturbed HF/6-311++G(2d,2p)//HF/6-311++G(2d,2p) level of theory. For the bending mode, shown in Fig. 3.13, the calculated dipole derivative is +0.879 a.u. directed along the y-axis. For the asymmetric stretch, shown in Fig. 3.14, the calculated dipole derivative is 3.793 a.u. directed along the z-axis. The difference between the calculated dipole derivatives is in good agreement with experimental results [15]. For the bending mode, the atomic polarization derivatives are counter to the larger magnitude atomic charge-transfer derivatives. In contrast, for the asymmetric stretching mode, both the atomic polarization derivatives and the atomic charge transfer derivatives contribute positively to the overall dipole derivative, making it much larger.

Table 3.4 Atomic contributions to the dipole moment derivatives of CO_2 (in a.u.) with respect to the bend and asymmetric stretch (AS) normal mode coordinates. Also shown are the atomic charge Q derivatives.[a,b,c]

Atom, Ω	$\mu^{Bend}(\Omega)_y$	$\mu_p{}^{Bend}(\Omega)_y$	$\mu_c{}^{Bend}(\Omega)_y$	$Q^{Bend}(\Omega)$
C1	+0.191	−0.741	+0.932	−0.012
O2	+0.342	−0.795	+1.137	+0.006
O3	+0.342	−0.795	+1.137	+0.006
Total	+0.876	−2.331	+3.207	+0.000

Atom, Ω	$\mu^{AS}(\Omega)_z$	$\mu_p{}^{AS}(\Omega)_z$	$\mu_c{}^{AS}(\Omega)_z$	$Q^{AS}(\Omega)$
C1	+1.655	+1.213	+0.442	−0.049
O2	+1.068	+0.780	+0.288	−0.508
O3	+1.055	+0.898	+0.157	+0.573
Total	+3.778	+2.891	+0.887	+0.016

[a] Calculated frequencies for the modes are 776 (bend) and 2551 (AS) cm^{-1}.
[b] Experimental frequencies for the modes are 667 (bend) and 2349 (AS) cm^{-1}.
[c] The experimentally measured ratio of the AS and bend dipole derivatives is 4.57 [15], compared to the calculated ratio of 4.31.

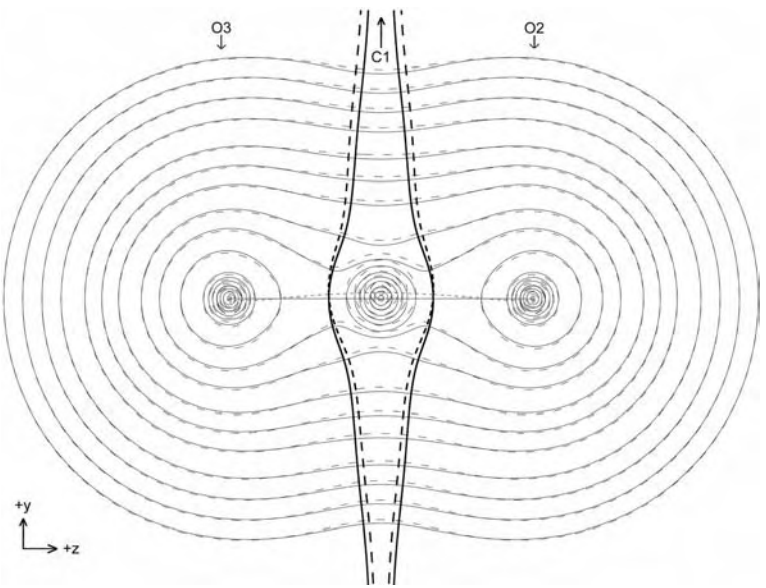


Fig. 3.13 Effect of finite ($\xi \approx 0.2$ a.u.) bending vibration on CO_2 in terms of electron density contours, interatomic surfaces (bold), and bond paths (semi-bold). Dotted lines correspond to bent geometry whereas solid lines correspond to equilibrium geometry.

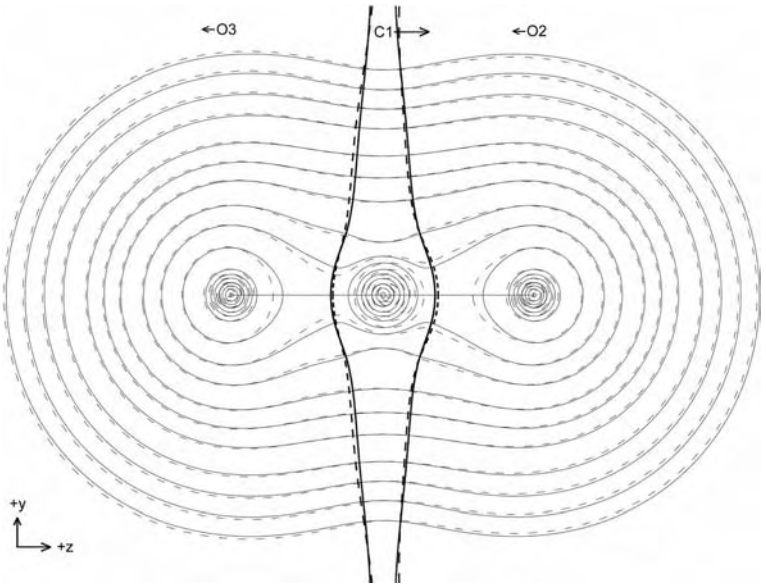


Fig. 3.14 Effect of finite ($\xi \sim 0.2$ a.u.) asymmetric stretching vibration on CO_2 in terms of electron density contours, interatomic surfaces (bold), and bond paths (semi-bold). Dotted lines correspond to asymmetrically stretched geometry whereas solid lines correspond to equilibrium geometry.

3.8
Atomic Nuclear Virial Energies

The electronic energy of an atom in a molecule, $E_e(\Omega)$, equivalent to minus the electronic kinetic energy of the atom, $T(\Omega)$, has a number of important characteristics [1], including:

- Well defined. The Hamiltonian (K) and Lagrangian (G) kinetic energy densities integrate to the same atomic electronic kinetic energy.
- Total energy additivity at stationary point geometries. The sum of the $E_e(\Omega)$ in a molecule at a stationary point geometry is equal to the total energy E of the molecule because of the molecular virial theorem.
- Transferability. The atomic electronic energies for different atoms – and their corresponding kinetic energy and virial field distributions – are similar in a manner which parallels the atomic charge distribution and which is consistent with experimentally based atomic and group additivity schemes.

- An atomic electronic virial theorem is satisfied: $E_e(\Omega) = -T(\Omega) = (1/2)V_e(\Omega)$, where $V_e(\Omega)$ is the atomic basin average of the virial of the electronic Ehrenfest force density plus the atomic surface average of the virial of the electronic pressure density, i.e. the atomic electronic potential energy contribution.
- Interpretable in terms of local, real space energy densities.

An open "problem" in QTAIM, however, is that for molecules at non-stationary point geometries the sum of the atomic electronic energies $E_e(\Omega)$ does not equal the total molecular energy E [1] because the forces on the nuclei (and hence the nuclear virial contribution to the energy) are not zero, i.e. $E \neq E_e$. This is easily seen from the molecular virial theorem [16], i.e., the hypervirial theorem for the full (electronic + nuclear) virial operator $\hat{\mathscr{V}}$:

$$\hat{\mathscr{V}} = \hat{\mathscr{V}}_e + \hat{\mathscr{V}}_n = (1/2)\sum_{i=1}^{N_e}[\hat{\mathbf{p}}_i \bullet \mathbf{r}_i + \mathbf{r}_i \bullet \hat{\mathbf{p}}_i] + (1/2)\sum_{A=1}^{N_a}[\hat{\mathbf{P}}_A \bullet \mathbf{R}_A + \mathbf{R}_A \bullet \hat{\mathbf{P}}_A] \tag{38}$$

$$\begin{aligned}(i/\hbar)\langle\psi|[\hat{\mathbf{H}}, \hat{\mathscr{V}}]|\psi\rangle &= (i/\hbar)\langle\psi|[E, \hat{\mathscr{V}}]|\psi\rangle \\ &= 2\langle\psi|\hat{\mathbf{T}}|\psi\rangle + \langle\psi|\hat{\mathbf{V}}_{ne}|\psi\rangle + \langle\psi|\hat{\mathbf{V}}_{ee}|\psi\rangle + \langle\psi|\hat{\mathbf{V}}_{nn}|\psi\rangle \\ &= -\sum_{A=1}^{N_a}\mathbf{R}_A \bullet \nabla_A E \\ &= 2T + V_{ne} + V_{ee} + V_{nn} = 2T + V = T + E = W\end{aligned} \tag{39}$$

Rearranging Eq. (39) gives:

$$E = -T + W = (1/2)(V + W) = E_e + W \tag{40}$$

The energy W may be called the "nuclear virial" contribution to the total molecular energy:

$$W = -\sum_{A=1}^{N_a}\mathbf{R}_A \bullet \nabla_A E = \sum_{A=1}^{N_a}\mathbf{R}_A \bullet \mathbf{G}_A \tag{41}$$

where $\mathbf{G}_A$ is the familiar *energy-gradient* based force on the nucleus A (which is, unfortunately, not equal to the Hellman–Feynman electrostatic $\mathbf{F}_A$ force on nucleus A in typical ab initio calculations since they typically violate the Hellman-Feynman electrostatic theorem):

$$\mathbf{G}_A = -\nabla_A E \tag{42}$$

In terms of atomic electronic energies $E_e(\Omega) = -T(\Omega)$ we have:

$$E = -T + W = \sum_{\Omega=1}^{N_a} -T(\Omega) + W = \sum_{\Omega=1}^{N_a} E_e(\Omega) + W \tag{43}$$

For molecular geometries at which $\mathbf{G}_A = 0$ for all nuclei (e.g. equilibrium and transition state geometries), W is zero and atomic electronic energies $E_e(\Omega)$ are additive to give the total molecular energy $E = E_e$.

For non-stationary point geometries this is not true and it is therefore important to be able to define a physically reasonable atomic contribution, $W(\Omega)$, to W so that the total energy additivity of atomic energies is preserved as a molecule vibrates, reacts, undergoes a vertical electronic transition or otherwise deviates from a stationary point geometry. In addition, even at equilibrium geometries a definition for $W(\Omega)$ is necessary if one wishes to calculate atomic contributions to molecule response properties defined in terms of energy derivatives with respect to nuclear coordinates (e.g. vibrational frequencies).

$$W = \sum_{\Omega=1}^{N_a} W(\Omega) \tag{44}$$

$$E(\Omega) = E_e(\Omega) + W(\Omega) = -T(\Omega) + W(\Omega) \tag{45}$$

$$E = \sum_{\Omega=1}^{N_a} E(\Omega) \tag{46}$$

In the limit of a stationary point geometry, where $W = 0$, each $W(\Omega)$ should separately equal zero so that the total atomic energy $E(\Omega)$ reduces to the atomic electronic energy $E_e(\Omega)$, which is already known to behave correctly at stationary point geometries:

$$\lim_{W \to 0}[W(\Omega)] = 0 \tag{47}$$

An obvious definition for $W(\Omega)$ which satisfies Eq. (47) would be:

$$\underline{W}_I(\Omega) = \mathbf{R}_\Omega \bullet \mathbf{G}_\Omega \tag{48}$$

where $\mathbf{G}_\Omega$ means the energy-gradient-based force on the nucleus of atom Ω and the subscript "I" is used to distinguish this definition from an alternative definition considered later.

As indicated by the underline, the definition for $\underline{W}_I(\Omega)$ in Eq. (48) is origin-dependent and is, therefore, not directly useful. It also has the physically unrea-

sonable characteristic that if the force on the nucleus of an atom is zero, its nuclear virial energy contribution is *necessarily* zero. To see that this is unphysical it is important to realize that the total force on a nucleus in a molecule is a molecular property, and therefore one to which each atom contributes [1]. This is evident from the Hellman–Feynman electrostatic theorem [16], discussed later in this section. Equation (48) does, however, provide a good starting point for defining a more physically reasonable $W_I(\Omega)$. Because of translation invariance [16], the sum **G** of the energy-gradient-based forces $\mathbf{G}_\Omega$ is zero, i.e. **G** is a "null" molecular property:

$$\mathbf{G} = \sum_{\Omega=1}^{N_a} \mathbf{G}_\Omega = 0 \tag{49}$$

Thus, the method of partitioning apparently origin-dependent atomic properties outlined in Sections 3.2 and 3.3 can be used to define an origin-independent $W_I(\Omega)$ based on the $\underline{W}_I(\Omega)$ given in Eq. (48):

$$\mathbf{G}_\Omega = \sum_{\Lambda=1}^{N_b(\Omega)} \mathbf{G}_\Omega(\Omega|\Lambda) \tag{50}$$

$$W_I(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] \bullet \mathbf{G}_\Omega(\Omega|\Lambda) \tag{51}$$

The force vector $\mathbf{G}_\Omega(\Omega|\Lambda)$ may be interpreted as a "bond force", or a bond contribution to the energy-gradient-based force on the nucleus of atom Ω. With this definition of an atomic nuclear virial energy, the corresponding definition of a total atomic energy $E_I(\Omega)$ in a molecule at an arbitrary geometry becomes:

$$E_I(\Omega) = -T(\Omega) + W_I(\Omega) = -T(\Omega) + \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] \bullet \mathbf{G}_\Omega(\Omega|\Lambda) \tag{52}$$

All of the terms in Eq. (52) are readily calculated from standard ab initio wavefunctions. For the additivity of Eq. (46) to be obtained in practice, however, the molecular virial theorem, Eq. (39), must be satisfied. Of course, satisfaction of the molecular virial theorem is also necessary at stationary point geometries for atomic energy additivity to be obtained in practice.

Unfortunately, the molecular virial theorem is *not* satisfied by wavefunctions from typical ab initio calculations [16]. It is, however, relatively straightforward to variationally improve any ab initio calculation to satisfy the molecular virial theorem using what this author calls "self-consistent virial scaling" (SCVS), which involves variationally scaling the electronic and nuclear coordinates of a wavefunction to satisfy the molecular virial theorem in a manner which is "self-consistent" with the ab initio method of choice. At the Hartree–Fock SCF level,

SCVS simply means optimizing the coordinate scaling parameter simultaneously with the MO coefficients – and geometry, if desired. The net result of SCVS is a legitimate ab initio wavefunction or first-order density matrix and molecular geometry that satisfies the molecular virial theorem.

Further information about the contributions to $W_I(\Omega)$ can be obtained by using the Hellman–Feynman electrostatic theorem [16], which is just the hypervirial theorem for nuclear momentum operator $\hat{\mathbf{P}}_A$:

$$\begin{aligned}(i/\hbar)\langle\psi|[\hat{H},\hat{\mathbf{P}}_A]|\psi\rangle &= (i/\hbar)\langle\psi|[E,\hat{\mathbf{P}}_A]|\psi\rangle = \langle\psi|(\nabla_A\hat{V}_{ne})|\psi\rangle + \nabla_A V_{nn}\\ &= -Z_A\left[\int d\mathbf{r}\rho(\mathbf{r})(\mathbf{r}-\mathbf{R}_A)|\mathbf{r}-\mathbf{R}_A|^{-3} + \sum_{B\neq A}^{N_a} Z_B(\mathbf{R}_B-\mathbf{R}_A)|\mathbf{R}_B-\mathbf{R}_A|^{-3}\right]\\ &= \mathbf{F}_A = -\nabla_A E = \mathbf{G}_A\end{aligned} \tag{53}$$

In words, the electrostatic force $\mathbf{F}_A$ on nucleus A is equal to the energy gradient based force $\mathbf{G}_A$ when the Hellman–Feynman electrostatic theorem is satisfied. Thus, if the Hellman–Feynman electrostatic theorem is satisfied, $W_I(\Omega)$ can be expressed as:

$$W_I(\Omega) = \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] \bullet \mathbf{F}_\Omega(\Omega|\Lambda) \tag{54}$$

where $\mathbf{F}_\Omega(\Omega|\Lambda)$ is the contribution from the bond between atoms Ω and Λ to the electrostatic force on the nucleus of atom Ω.

In practice, unfortunately, the Hellman–Feynman electrostatic theorem is not satisfied in typical ab initio calculations [16], and improving typical ab initio calculations to produce legitimate wavefunctions which satisfy the Hellman–Feynman electrostatic theorem is somewhat more difficult than for the molecular virial theorem.

Practical difficulties aside, the Hellman–Feynman electrostatic theorem does provide an alternative way to define a $W(\Omega)$ which should be considered. Combining Eqs (41) and (53) the total nuclear virial energy W can be expressed as:

$$\begin{aligned}W &= \sum_{A=1}^{N_a} \mathbf{R}_A \bullet \mathbf{G}_A = \sum_{A=1}^{N_a} \mathbf{R}_A \bullet \mathbf{F}_A\\ &= \sum_{A=1}^{N_a} Z_A\mathbf{R}_A \bullet \left[-\int d\mathbf{r}\rho(\mathbf{r})(\mathbf{r}-\mathbf{R}_A)|\mathbf{r}-\mathbf{R}_A|^{-3}\right.\\ &\qquad\left. + \sum_{B\neq A}^{N_a} Z_B(\mathbf{R}_B-\mathbf{R}_A)|\mathbf{R}_B-\mathbf{R}_A|^{-3}\right]\end{aligned} \tag{55}$$

The contribution, $\mathbf{F}_A(\Omega)$, of atom Ω to the electrostatic force $\mathbf{F}_A$ on nucleus A is easily determined (without the Hellman–Feynman electrostatic theorem, the contribution of Ω to $\mathbf{G}_A$ is not easily determined):

$$\mathbf{F}_A(\Omega) = -\int_\Omega d\mathbf{r}\rho(\mathbf{r})(\mathbf{r}-\mathbf{R}_A)|\mathbf{r}-\mathbf{R}_A|^{-3} + Z_A Z_\Omega(\mathbf{R}_\Omega - \mathbf{R}_A)|\mathbf{R}_\Omega - \mathbf{R}_A|^{-3} \quad (56)$$

Thus, W can be expressed as:

$$W = \sum_{A=1}^{N_a} \mathbf{R}_A \bullet \mathbf{F}_A = \sum_{A=1}^{N_a}\sum_{\Omega=1}^{N_a} \mathbf{R}_A \bullet \mathbf{F}_A(\Omega) \quad (57)$$

Using this expression, an origin-dependent atomic contribution, $\underline{W}_{II}(\Omega)$, to W can be defined:

$$\underline{W}_{II}(\Omega) = \sum_{A=1}^{N_a} \mathbf{R}_A \bullet \mathbf{F}_A(\Omega) = \sum_{A=1}^{N_a} (\mathbf{R}_A - \mathbf{R}_\Omega) \bullet \mathbf{F}_A(\Omega) + \mathbf{R}_\Omega \bullet \sum_{A=1}^{N_a} \mathbf{F}_A(\Omega)$$

$$= \sum_{A=1}^{N_a} (\mathbf{R}_A - \mathbf{R}_\Omega) \bullet \mathbf{F}_A(\Omega) + \mathbf{R}_\Omega \bullet \mathbf{F}_{Tot}(\Omega) \quad (58)$$

where $\mathbf{F}_{Tot}(\Omega)$ is defined as the sum of the electrostatic forces that the charge distribution (electron and nuclear) of atom Ω exerts on all the nuclei. When the Hellman–Feynman electrostatic theorem is satisfied, the total $\mathbf{F}_{Tot}$ is a "null" molecular property, because of translational invariance [16]:

$$\mathbf{F}_{Tot} = \sum_{\Omega=1}^{N_a} \mathbf{F}_{Tot}(\Omega) = 0 \quad (59)$$

An origin-independent $W_{II}(\Omega)$ can therefore be expressed according to Sections 3.2 and 3.3 as:

$$W_{II}(\Omega) = \sum_{A=1}^{N_a} (\mathbf{R}_A - \mathbf{R}_\Omega) \bullet \mathbf{F}_A(\Omega) + \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] \bullet \mathbf{F}_{Tot}(\Omega|\Lambda) \quad (60)$$

and the corresponding total atomic energy $E_{II}(\Omega)$ is given by:

$$E_{II}(\Omega) = -T(\Omega) + W_{II}(\Omega) \quad (61)$$

To summarize the meaning of the terms in Eq. (60), $\mathbf{F}_A(\Omega)$ is the electrostatic force that the charge distribution (electronic and nuclear) of atom Ω exerts on the

nucleus A while $\mathbf{F}_{\mathrm{Tot}}(\Omega|\Lambda)$ is the contribution from the (directed) bond between atoms Ω and Λ to the electrostatic force that the charge distribution of atom Ω exerts on the *all* the nuclei in the molecule.

The two definitions for $\mathrm{W}(\Omega)$ considered here, $\mathrm{W_I}(\Omega)$ in Eq. (51) and $\mathrm{W_{II}}(\Omega)$ in Eq. (60), both employ the strategy of Sections 3.2 and 3.3 to avoid apparent origin dependence, but $\mathrm{W_I}(\Omega)$ and $\mathrm{W_{II}}(\Omega)$ are otherwise different. $\mathrm{W_I}(\Omega)$ has the advantage that it is relatively easily and reliably calculated as long as the molecular virial theorem is satisfied whereas $\mathrm{W_{II}}(\Omega)$ requires satisfaction of both the molecular virial theorem and the more difficult Hellman–Feynman theorem for its application. $\mathrm{W_{II}}(\Omega)$, however, has the important advantage that its definition is more consistent with the definition of other atomic properties, and provides a means for a richer interpretation directly in terms of the charge distribution. Both definitions, and perhaps others not considered here, should be investigated.

3.9
Atomic Contributions to Induced Electronic Magnetic Dipole Moments

The electronic magnetic dipole moment, $\mathbf{m}$, of a closed-shell molecule in a uniform magnetic field $\mathbf{B}$ is given by [13, 17]:

$$\mathbf{m} = (1/2c)\int[\mathbf{r}-\mathbf{R}_0]\times\mathbf{J}(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{62}$$

where $\mathbf{J}(\mathbf{r})$ is the electronic current density at a point $\mathbf{r}$ in real space and $\mathbf{R}_0$ is the arbitrary origin of the molecular coordinate system. Because the net current $\langle\mathbf{J}\rangle$ is zero for the molecule, $\mathbf{m}$ is independent of $\mathbf{R}_0$.

Like the electric dipole moment $\boldsymbol{\mu}$, the magnetic dipole moment $\mathbf{m}$ can be expressed in terms of atomic magnetic polarization contributions, $\mathbf{m}_{\mathrm{p}}(\Omega)$, and *origin-dependent* atomic "net current" contributions, $\boldsymbol{m}_{\mathrm{c}}(\Omega)$, as follows:

$$\begin{aligned}
\mathbf{m} &= (1/2c)\sum_{\Omega=1}^{\mathrm{N_a}}\left\{\int_\Omega[\mathbf{r}-\mathbf{R}_0]\times\mathbf{J}(\mathbf{r})\,\mathrm{d}\mathbf{r}\right\}\\
&= (1/2c)\sum_{\Omega=1}^{\mathrm{N_a}}\left\{\int_\Omega[\mathbf{r}-\mathbf{R}_\Omega]\times\mathbf{J}(\mathbf{r})\,\mathrm{d}\mathbf{r}+(\mathbf{R}_\Omega-\mathbf{R}_0)\times\left[\int_\Omega\mathbf{J}(\mathbf{r})\,\mathrm{d}\mathbf{r}\right]\right\}\\
&= (1/2c)\sum_{\Omega=1}^{\mathrm{N_a}}\left\{\int_\Omega[\mathbf{r}-\mathbf{R}_\Omega]\times\mathbf{J}(\mathbf{r})\,\mathrm{d}\mathbf{r}+[\mathbf{R}_\Omega-\mathbf{R}_0]\times\mathbf{J}(\Omega)\right\}\\
&= \sum_{\Omega=1}^{\mathrm{N_a}}\{\mathbf{m}_{\mathrm{p}}(\Omega)+\underline{\mathbf{m}}_{\mathrm{c}}(\Omega)\}=\sum_{\Omega=1}^{\mathrm{N_a}}\underline{\mathbf{m}}(\Omega)=\mathbf{m}_{\mathrm{p}}+\mathbf{m}_{\mathrm{c}}
\end{aligned} \tag{63}$$

where $\mathbf{J}(\Omega)$ is the net current of atom Ω. As in earlier sections, underlines are used to emphasize that $\underline{\mathbf{m}}_c(A)$, and hence $\underline{\mathbf{m}}(\Omega)$, is origin-dependent. Using Eq. (7), $\underline{\mathbf{m}}(\Omega)$ can be expressed in terms of "bond current" contributions $\mathbf{J}(\Omega|\Lambda)$ as:

$$\underline{\mathbf{m}}_c(\Omega) = (1/2c)[\mathbf{R}_\Omega - \mathbf{R}_0] \times \mathbf{J}(\Omega) = (1/2c)[\mathbf{R}_\Omega - \mathbf{R}_0] \times \sum_{\Lambda=1}^{N_b(\Omega)} \mathbf{J}(\Omega|\Lambda) \tag{64}$$

Using Eq. (8), a corresponding origin-independent expression, $\mathbf{m}_c(\Omega)$, can be defined:

$$\mathbf{m}_c(\Omega) = (1/2c) \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|\Lambda)] \times \mathbf{J}(\Omega|\Lambda) \tag{65}$$

where $\mathbf{J}(\Omega|\Lambda)$ is the contribution to the net current of atom from the bond from atom Ω to atom Λ.

Note that for each term $\mathbf{R}_b(\Omega|\Lambda) \times \mathbf{J}(\Omega|\Lambda)$ for atom Ω, there is a corresponding term for atom Λ which cancels it, because of Eq. (8), and therefore:

$$\sum_{\Omega=1}^{N_a} \mathbf{m}_c(\Omega) = (1/2c) \sum_{\Omega=1}^{N_a} \sum_{\Lambda=1}^{N_b(\Omega)} [\mathbf{R}_\Omega - \mathbf{R}_b(\Omega|B)] \times \mathbf{J}(\Omega|B)$$

$$= (1/2c) \sum_{\Omega=1}^{N_a} \mathbf{R}_\Omega \times \mathbf{J}(\Omega) = \sum_{\Omega=1}^{N_a} \underline{\mathbf{m}}_c(\Omega) \tag{66}$$

$$\mathbf{m} = \sum_{\Omega=1}^{N_a} \mathbf{m}(\Omega) = \sum_{\Omega=1}^{N_a} \{\mathbf{m}_p(\Omega) + \mathbf{m}_c(\Omega)\} = \mathbf{m}_p + \mathbf{m}_c \tag{67}$$

The meaning of $\mathbf{m}_p(\Omega)$ and $\mathbf{m}_c(\Omega)$ is essentially that the former describes the magnetic moment arising from current flowing within the atomic basin whereas the latter describes the magnetic moment from current flowing between atomic basins. When a current distribution is highly localized, $\mathbf{m}_p(\Omega)$ is larger than $\mathbf{m}_c(\Omega)$. When a current distribution is highly delocalized, the opposite is true.

Note that all of the quantities in the expressions for $\mathbf{m}_p(\Omega)$ and $\mathbf{m}_c(\Omega)$ are uniquely determined by the molecular charge distribution and current density distributions.

In the absence of an external magnetic field $\mathbf{B}$, the current density $\mathbf{J}(\mathbf{r})$ vanishes for closed-shell molecules [13]. For open-shell molecules, $\mathbf{J}(r)$ does not necessarily vanish even when $\mathbf{B} = 0$ and in such circumstances the corresponding permanent molecular electronic magnetic moment can be expressed in terms of the atomic contributions $\mathbf{m}_p(\Omega)$ and $\mathbf{m}_c(\Omega)$ described here.

3.10
Atomic Contributions to Magnetizabilities of Closed-Shell Molecules

Analogous to the electric polarizability tensor α, the magnetizability tensor of a molecule, χ, is the gradient of the molecular electronic magnetic dipole moment, $\mathbf{m}$, with respect to a uniform magnetic field, $\mathbf{B}$, in the limit of zero field strength [13]:

$$\chi = [\nabla_{\mathbf{B}}\mathbf{m}]_{B=0} = [\mathbf{i}(\partial\mathbf{m}/\partial B_x) + \mathbf{j}(\partial\mathbf{m}/\partial B_y) + \mathbf{k}(\partial\mathbf{m}/\partial B_z)]_{B=0} = \mathbf{m}^{\mathbf{B}} \tag{68}$$

where the notation $t^{\mathbf{B}}$ is used to indicate the gradient of the term t with respect to $\mathbf{B}$, evaluated at $\mathbf{B} = 0$.

From Eq. (62), χ is given by:

$$\chi = (1/2c)\int(\mathbf{r} - \mathbf{R}_0) \times \mathbf{J}^{\mathbf{B}}(\mathbf{r})\,d\mathbf{r} \tag{69}$$

where $\mathbf{J}^{\mathbf{B}}(\mathbf{r})$ is the gradient of the electronic current density $\mathbf{J}(\mathbf{r})$ with respect to $\mathbf{B}$, in the limit of zero field strength:

$$\mathbf{J}^{\mathbf{B}}(\mathbf{r}) = [\nabla_{\mathbf{B}}\mathbf{J}(\mathbf{r})]_{B=0} = [\mathbf{i}(\partial\mathbf{J}(\mathbf{r})/\partial B_x) + \mathbf{j}(\partial\mathbf{J}(\mathbf{r})/\partial B_y) + \mathbf{k}(\partial\mathbf{J}(\mathbf{r})/\partial B_z)]_{B=0} \tag{70}$$

The first-order current density $\mathbf{J}^{(1)}(\mathbf{r})$ induced by $\mathbf{B}$ is:

$$\mathbf{J}^{(1)}(\mathbf{r}) = \mathbf{J}^{\mathbf{B}}(\mathbf{r}) \bullet \mathbf{B} \tag{71}$$

Methods for calculating relatively accurate $\mathbf{J}^{(1)}$ distributions and their dependent properties, for example magnetizability tensors discussed in this section and NMR shielding tensors [18], were developed by Keith and Bader [5, 19]. Thorough and correct displays and analyses of $\mathbf{J}^{(1)}$ distributions were presented by Keith and Bader [20] based on methods developed elsewhere [5, 19], along with pioneering work by Gomes [21].

In terms of atomic contributions to χ, from Eq. (67) we have:

$$\begin{aligned}\chi &= \sum_{\Omega=1}^{N_a} \mathbf{m}^{\mathbf{B}}(\Omega) = \sum_{\Omega=1}^{N_a}\{\mathbf{m}_p{}^{\mathbf{B}}(\Omega) + \mathbf{m}_c{}^{\mathbf{B}}(\Omega)\} \\ &= \sum_{\Omega=1}^{N_a} \chi_p(\Omega) + \chi_c(\Omega) = \sum_{\Omega=1}^{N_a} \chi(\Omega)\end{aligned} \tag{72}$$

where $\chi_p(\Omega)$ and $\chi_c(\Omega)$ are given in terms of the current density by:

$$\boldsymbol{\chi}_{\mathrm{p}}(\Omega) = (1/2\mathrm{c}) \int_{\Omega} (\mathbf{r} - \mathbf{R}_{\Omega}) \times \mathbf{J}^{\mathbf{B}}(\mathbf{r})\, \mathrm{d}\mathbf{r} \tag{73}$$

$$\boldsymbol{\chi}_{\mathrm{c}}(\Omega) = (1/2\mathrm{c}) \sum_{\Lambda=1}^{\mathrm{N_b}(\Omega)} \{[\mathbf{R}_{\Omega} - \mathbf{R}_{\mathrm{b}}(\Omega, \Lambda)] \times \mathbf{J}^{\mathbf{B}}(\mathbf{r})(\Omega|\Lambda) \tag{74}$$

Note that, unlike the atomic electric polarizability tensors $\boldsymbol{\alpha}(\Omega)$, there are no surface derivative contributions to $\boldsymbol{\chi}(\Omega)$, because $[\nabla_{\mathbf{B}}\rho(\mathbf{r})]_{\mathrm{B}=0} = 0$, i.e. the first-order correction to the electron density from a magnetic field perturbation vanishes because the magnetic field perturbation term in the Hamiltonian (and hence the first-order perturbed wavefunction) are purely imaginary functions [13].

Using these formulas, Bader and Keith [6, 7] studied the atomic contributions to the magnetizability tensors of several series of molecules and the reader is referred to those papers for details. Especially noteworthy from these studies is that for the normal alkane series, the QTAIM isotropic magnetic susceptibilities of the methylene and methyl groups matched the transferable behavior of other properties for these groups and the magnetic susceptibility of the transferable methyl and methylene groups of the series matched those of Pascal [22]. Also noteworthy was that for the benzene molecule it was found that the threefold increase in magnetic susceptibility for a field applied perpendicular to the plane of the ring, compared with that for a field applied parallel to the plane of the ring, was largely because of the $\boldsymbol{\chi}_{\mathrm{c}}(\Omega)$ contributions from the carbon atoms, i.e. the flow of current between the carbon atoms, thus providing a fundamental, physical justification for the famous ring current model of benzene. The significance of this study should not be overlooked. The validity of the ring-current model as an explanation for the magnetic response properties of benzene and other aromatic systems has been debated for decades, usually in terms of orbital models [23]. Keith and Bader analyzed the validity of the ring current model in the only physically reasonable way possible – by actually identifying the atoms in benzene and quantifying the contribution to the magnetizability tensor in terms of the flow of total, physical current within the atoms and between the atoms.

As an additional example of this kind, Table 3.5 shows the principal components of $\boldsymbol{\chi}_{\mathrm{p}}(\Omega)$, $\boldsymbol{\chi}_{\mathrm{c}}(\Omega)$, and $\boldsymbol{\chi}(\Omega)$ for the symmetrically unique atoms of naphthalene (Figs 3.15 and 3.16), calculated at coupled–perturbed HF/6-311++G(2d,2p)// HF/6-311++G(2d,2p). Also shown are the isotropically averaged contributions $\boldsymbol{\chi}_{\mathrm{iso}}$ and the contributions to the major anisotropy $\boldsymbol{\chi}_{\mathrm{aniso}}$. From these results it is apparent that the vast majority of the anisotropy of the magnetic susceptibility in naphthalene is because of the $\boldsymbol{\chi}_{\mathrm{c}}(\Omega)$ contributions from the carbon atoms, which is again consistent with the ring-current model. Note that the $\boldsymbol{\chi}_{\mathrm{c}}(\Omega)_{zz}$ contribution for each of the two fused carbon atoms, C1 and C6, is approximately 50% larger than that for the other carbons, because the fused carbon atoms have three C–C bonds instead of two and it is the flow of current across the C–C surfaces which is responsible for the large $\boldsymbol{\chi}_{\mathrm{c}}(\Omega)_{zz}$ for the carbon atoms. These numbers are reflected in the current displays shown in Figs 3.15 and 3.16.

Table 3.5 Atomic and bond contributions (in cgs-ppm) to the magnetizability tensor of naphthalene (B/S).

Atom, Ω	$\chi(\Omega)_{zz}$	$\chi_p(\Omega)_{zz}$	$\chi_c(\Omega)_{zz}$	$\chi_c(\Omega\|\Lambda_1)_{zz}\Lambda_1$	$\chi_c(\Omega\|\Lambda_2)_{zz}\Lambda_2$	$\chi_c(\Omega\|\Lambda_3)_{zz}\Lambda_3$
C1	−21.020	−2.850	−18.171	−7.778 C6	−5.196 C4	−5.196 C9
C4	−15.183	−3.145	−12.038	−5.514 C5	−5.139 C6	−1.385 H13
C7	−16.257	−3.567	−12.689	−5.479 C8	−5.865 C10	−1.345 H17
H11	−2.079	−1.294	−0.785	−0.785 C10		
H12	−2.083	−1.273	−0.811	−0.811 C9		
Total	−184.449	−42.812	−141.637			

Atom, Ω	$\chi(\Omega)_{yy}$	$\chi_p(\Omega)_{yy}$	$\chi_c(\Omega)_{yy}$	Atom, Ω	$\chi(\Omega)_{xx}$	$\chi_p(\Omega)_{xx}$	$\chi_c(\Omega)_{xx}$
C1	−3.091	−0.978	−2.113	C1	−3.622	−1.241	−2.381
C4	−3.786	−1.182	−2.604	C4	−5.336	−3.351	−1.985
C7	−5.184	−2.761	−2.423	C7	−4.006	−1.910	−2.095
H11	−1.827	−1.316	−0.511	H11	−1.425	−1.274	−0.151
H12	−1.194	−1.194	0.000	H12	−1.978	−1.322	−0.656
Total	−54.143	−27.765	−26.378	Total	−58.223	−33.913	−24.310

Atom, Ω	$\chi(\Omega)_{iso}$	$\chi_p(\Omega)_{iso}$	$\chi_c(\Omega)_{iso}$	Atom, Ω	$\chi(\Omega)_{aniso}$	$\chi_p(\Omega)_{aniso}$	$\chi_c(\Omega)_{aniso}$
C1	−9.244	−1.689	−7.555	C1	−17.664	−1.740	−15.924
C4	−8.102	−2.559	−5.542	C4	−10.622	−0.878	−9.744
C7	−8.482	−2.746	−5.736	C7	−11.662	−1.232	−10.43
H11	−1.777	−1.295	−0.483	H11	−0.452	+0.001	−0.454
H12	−1.752	−1.263	−0.489	H12	−0.498	−0.014	−0.483
Total[a]	−98.938	−34.830	−64.108	Total	−128.267	−11.973	−116.293

[a] An experimental value for the isotropic magnetizability of naphthalene is −91.6 cgs-ppm [24].

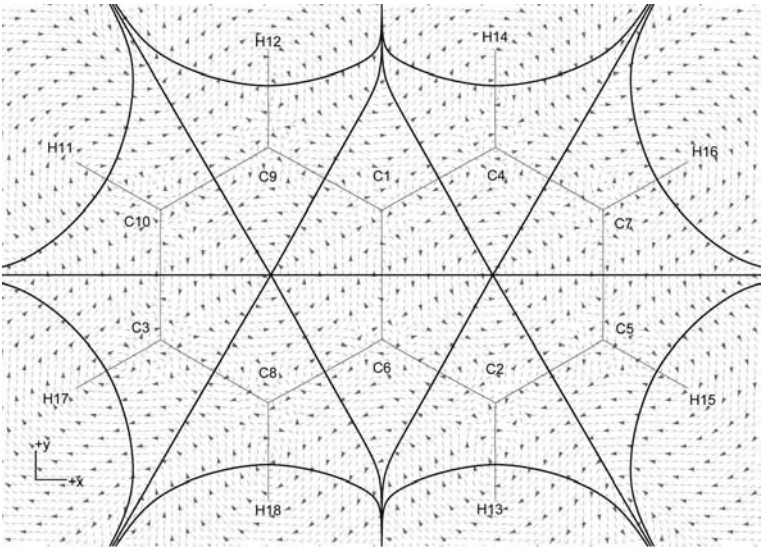


Fig. 3.15 Current density trajectories induced in the nuclear plane of naphthalene by a magnetic field applied along the +Z axis, i.e. coming out of the page. Arrows indicate direction of current flow. Counterclockwise flow is paramagnetic whereas clockwise flow is diamagnetic. Interatomic surfaces (bold) and bond paths (semi-bold) are also shown.

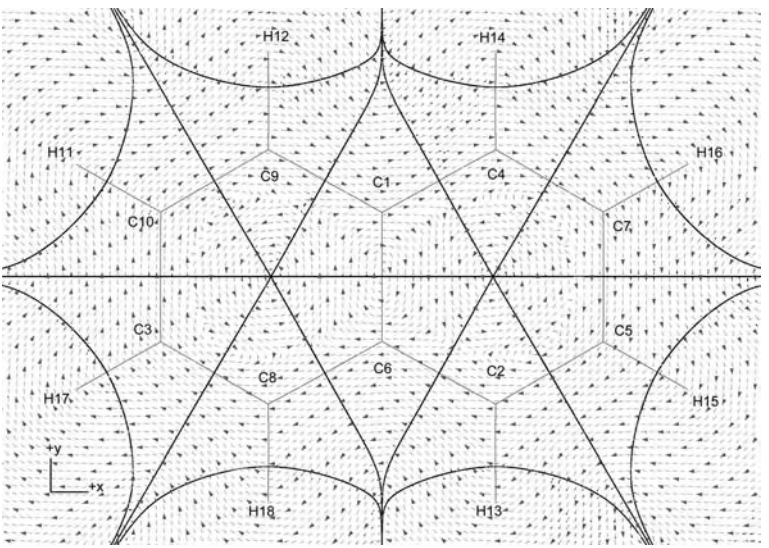


Fig. 3.16 Current density trajectories which intersect a plane ~1.0 a.u. above the nuclear plane of naphthalene and induced by a magnetic field applied along +Z axis, i.e. coming out of the page. Arrows indicate direction of current flow. Counterclockwise flow is paramagnetic whereas clockwise flow is diamagnetic. Interatomic surfaces (bold) and bond paths (semi-bold) in the nuclear plane are also shown.

References

1 R.F.W. Bader, *Atoms in Molecules: a Quantum Theory*, Oxford University Press, Oxford UK, **1990**.
2 R.F.W. Bader, A. Larouche, C. Gatti, M.T. Carroll, P.J. MacDougall and K.B. Wiberg, *J. Chem. Phys*, **1987**, 87, 1142.
3 K.E. Laidig and R.F.W. Bader, *J. Chem. Phys*, **1990**, 93, 7213.
4 R.F.W. Bader, T.A. Keith, K.M. Gough and K.E. Laidig, *Mol. Phys.*, **1992**, 75, 1167.
5 T.A. Keith and R.F.W. Bader, *Chem. Phys. Lett.*, **1992**, 194, 1.
6 R.F.W. Bader and T.A. Keith, *J. Chem. Phys.*, **1993**, 99, 3683.
7 R.F.W. Bader and T.A. Keith, *Int. J. Quant. Chem.*, **1996**, 60, 373.
8 R.F.W. Bader, D. Bayles and G.L. Heard, *J. Chem. Phys.*, **2000**, 112, 10095.
9 R.F.W. Bader and C.F. Matta, Int. J. Quant. Chem., 2001, 85, 592.
10 R.F.W. Bader, *Mol. Phys.*, **2002**, 100, 3333.
11 LAPACK – Linear Algebra PACKage, http://www.netlib.org/lapack
12 L.-T. Ching, W. Tam, S.H. Stevenson, G. Meredith, G. Rikken and S.R. Marder, *J. Phys. Chem.*, **1991**, 95, 10631.
13 P.W. Atkins, *Molecular Quantum Mechanics – Second Edition*, Oxford University Press, Oxford UK, **1983**.
14 E.B. Wilson Jr., J.C. Decius and P.C. Cross, *Molecular Vibrations*, Dover Publications, Inc, New York, **1955**.
15 Y. Yamaguchi, M. Frisch, J. Gaw, H.F. Schaefer III and J.S. Binkley, *J. Chem. Phys.*, **1986**, 84, 2262.
16 S.T. Epstein, *The Variation Method in Quantum Chemistry*, Academic Press, NY, **1974**.
17 C.J. Jameson and A.D. Buckingham, *J. Chem. Phys.*, **1980**, 73, 5684.
18 T.A. Keith and R.F.W. Bader, *Can. J. Chem.*, **1996**, 74, 185.
19 T.A. Keith and R.F.W. Bader, *Chem. Phys. Lett.*, **1993**, 210, 223.
20 T.A. Keith and R.F.W. Bader, *J. Chem. Phys.*, **1993**, 99, 3669.
21 J.A.N.F. Gomes, *Phys. Rev. A*, **1983**, 28, 559.
22 P. Pascal, Ann. Chim. Phys., **1910**, 19, 5.
23 P. Lazzeretti, *Prog. in Nuc. Mag. Res. Spec.*, **2000**, 36, 1.
24 *CRC Handbook of Chemistry and Physics*, 82nd edn, **2001–2002**.

4
QTAIM Analysis of Raman Scattering Intensities: Insights into the Relationship Between Molecular Structure and Electronic Charge Flow

Kathleen M. Gough, Richard Dawes, Jason R. Dwyer, and Tammy L. Welshman

4.1
Introduction

The investigation of the interplay between molecular structure and molecular properties is one that still captivates the imagination. The chemist today relies on computational modeling to gain insight into known properties and to help predict the behavior of systems for which experimental data are lacking. This two-way exchange epitomizes the conversation between experimentalists and theoreticians and shows us the way forward in developing our science. New theories breathe new life and meaning into existing experimental techniques and lead to the discovery of new truths and insight. Nature has all the answers and we have but to find them. In this chapter we will outline one such conversation – that between Raman spectroscopy and the quantum theory of atoms in molecules (QTAIM). Absolute Raman scattering intensities are difficult to measure yet they contain much valuable information. As such, they are intrinsically interesting, but they also provide an excellent framework on which to compose another theoretical–experimental conversation. QTAIM brings a unique and critical element to this enterprise; through this analysis, Raman spectroscopy is cast as an extremely sensitive probe of the structural origins of molecular function. We will discover that despite experimental and computational challenges, this dual-pronged approach yields great rewards.

We have seen in previous chapters the power of real-space analysis and the transferability of some functional group properties but now we will examine nontransferable properties. Nontransferable properties are those molecular properties which cannot be partitioned into functional groups independent of the molecular setting. It must be emphasized that the failure of some molecular properties to partition into neat functional group categories is no failure of QTAIM, but is instead a reflection of nature. The power of QTAIM in these instances is to analyze the detailed electronic distribution using the convenience and familiarity of real-space rather than an unwieldy multidimensional wavefunction. The electron dis-

tributions of molecules with "fixed" nuclei have been well characterized by use of QTAIM [1] and X-ray crystallographic experiments [2]. We now seek the response of the electron distribution to an infinitesimal structural perturbation in the presence of a field that can easily be related to the experimental Raman intensity of a given vibration.

Chemists traditionally think of molecular vibrations according to two different paradigms – normal mode analysis and functional group identification. The former is well understood [3, 4] but the quantity and quality of detailed experimental information required to obtain a reliable, experimentally refined harmonic force field ensures that this approach is usually restricted to molecules with no more than 20 atoms [5, 6]. The latter is derived from the nature of molecular vibrations, because shared functionality often produces similar characteristic vibrational modes (e.g. CH stretching modes lie between 2700 and 3200 cm^{-1}; C=O stretches will be found between 1600 and 1900 cm^{-1}).

The vibrational frequency is based on mechanical properties (anharmonically oscillating atomic masses) whereas the intensity is a function of the change in permanent (infrared) or induced (Raman) dipole moment. Normal mode analyses of larger molecules [6] reveal that vibrations assigned to a particular functional group always involve some displacement of many other nuclei. Thus, the variability in the frequency of modes assigned to specific functional groups is because a molecular vibration is always a property of the entire molecule. We continue to use functional group assignment because of its simplicity and utility, while recognizing the limitations.

Analysis of vibrational intensities has often been approximated in terms of changes in the local dipole moment (permanent or induced) associated with a change in functional group geometry. The same limitations and caveats – that functional groups are useful organizing principles but not rigorously transferable – apply, perhaps even more strongly. In this chapter, we will present a brief introduction to the problem of Raman scattering intensities, a detailed description of our theoretical approach to its solution, and a discussion of the insight we have achieved by means of QTAIM analysis. The resulting insights go beyond an explanation of Raman intensities and lead to improved understanding of the very underpinnings of molecular properties and charge flow. This is accompanied by some worked examples and suggested exercises for the reader.

4.2 Background to the Problem

Experimental measurements of absolute Raman trace scattering cross-sections provide a direct physical link to the redistribution of charge in a vibrating molecule. As such, they are an excellent probe for exploring the behavior of molecules in electric fields. The magnitude of the scattering cross-section, also called intensity, depends on the magnitude of the change in the molecular polarizability ten-

sor, α, during a vibration, where α is a measure of the magnitude of the dipole moment (μ) induced in a molecule when an electric field (E) is applied:

$$\mu = \alpha E \tag{1}$$

The derivative of the polarizability with respect to some normal mode of vibration (q) is denoted $\partial\alpha/\partial$q. Our original interest in this problem derived from the failure of the bond polarizability model [7], an early, extant description of Raman intensities. This model simplifies the problem by considering α to be a sum of independent bond polarizability ellipsoids, directed along each bond, in much the same way that the molecular dipole moment is sometimes approximated as a sum of bond dipoles. Parameters for the bond polarizability model were constructed from experimental data, primarily on alkanes, and some limited transferability between molecules was reported [7]. One of the most commonly used assumptions was that the derivative for the CH stretch vibrations was unchanged for a wide variety of hydrocarbons, including saturated alkanes, alkenes, alkynes, and aromatic compounds. Our experimental studies show that this approximation is not valid, even for relatively simple alkanes [6, 8–13]. It is moderately valid when the average intensity of a few small molecules is considered; it fails completely, however, when individual bonds are examined by selective deuteration. The model fails because the derivative, and hence the observed scattering intensity, arises from a change in the mean *molecular* polarizability, in turn dependent on charge redistribution throughout the entire molecule, not just within one bond.

4.2.1 Conceptual Approach to a Solution

Our guiding principle has been the conversation between experiment and theory. Raman trace scattering intensities provide information about the change in molecular charge distribution during a vibration. Ab initio calculations yield wavefunctions that, through QTAIM, let us examine the atom-by-atom response of the charge density to the applied field. Our early work was directed toward identification of the variability in polarizability derivatives for a series of simple alkanes. Subsequent theoretical modeling revealed the presence of patterns that enabled us to interpret and predict the behavior of much larger systems.

4.2.1.1 Experimental Measurement of Raman Scattering Intensities

The process of Raman scattering is well understood and is described in many standard texts [14]. According to the Placzek polarizability theory [15] the differential Raman scattering cross-section of a fundamental vibrational band ($\partial\sigma/\partial\Omega$, where Ω = solid angle in this expression) depends on the square of the derivative of the invariants of the molecular polarizability ($\bar{\alpha}$ and γ):

$$\left(\frac{\partial\sigma}{\partial\Omega}\right)_i = \frac{\pi^2}{90\varepsilon_0{}^2}\,\frac{(\bar{\nu}_0 - \bar{\nu}_i)^4}{1 - e^{-\bar{\nu}_i hc/kT}}\,g_i\left[45\left(\frac{\partial\bar{\alpha}}{\partial q_i}\right)^2 + 7\left(\frac{\partial\gamma}{\partial q_i}\right)^2\right] \tag{2}$$

where $\bar{\nu}_0$ and $\bar{\nu}_i$ are the wavenumbers of incident light and the vibration, respectively, g_i is the degeneracy of the ith vibrational mode, and h, k, T, and ε_0 have the usual meaning (details are given in our experimental papers [6, 8–13]). This theory assumes that mechanical and electrical anharmonicity are not significant.

The first indication of the inadequacy of transferability appeared during a study of the correlation between bond length and CH stretch frequency for a series of perdeuterated alkanes [16]. For cyclohexane-d_{11}, the normal mode for the CH stretch is essentially the isolated stretch of the lone CH bond. If transferability were correct, the intensity ratio of the bands from the axial (CH_{ax}) and equatorial (CH_{eq}) stretch vibrations would be 1, given that in the gas phase the molecule would exist as a 50:50 mixture of the two possible CH orientations. In fact, it is 0.7 (Fig. 4.1) [16].

Since that paper appeared, the absolute Raman trace scattering intensities have been reported for ethane [8], propane [9], *n*-butane [10], cyclohexane [11], ethene [12], ethyne [12], and, most recently, bicyclo-[1.1.1]-pentane [13]. For the purposes of the investigation, we restricted our attention to the trace scattering ($\partial\bar{\alpha}/\partial q$), that is, the change in the mean molecular polarizability that arises from the totally symmetric vibrations, where:

$$\frac{\partial\bar{\alpha}}{\partial q} = \frac{1}{3}\left(\frac{\partial\alpha_{xx}}{\partial q} + \frac{\partial\alpha_{yy}}{\partial q} + \frac{\partial\alpha_{zz}}{\partial q}\right) \tag{3}$$

Experimental error is reduced, because these modes produce a strong, narrow Q branch, due to more restrictive rotational selection rules. In addition, the trace

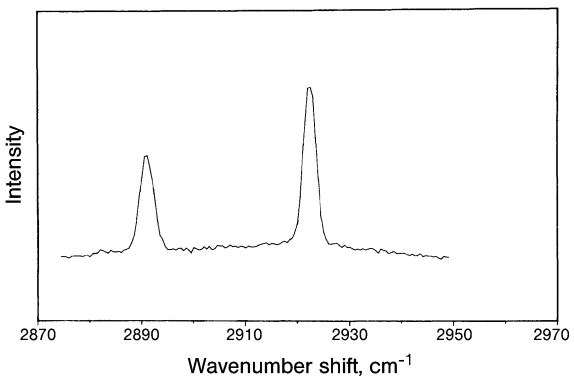


Fig. 4.1 Isolated CH stretching bands in the Raman spectrum of gaseous cyclohexane. Randomly protonated C_6HD_{11}, showing the hydrogen impurity in the perdeuterated sample. Reproduced, with permission, from *J. Chem. Phys.* **1984**, *81*, 5352–5361, American Institute of Physics.

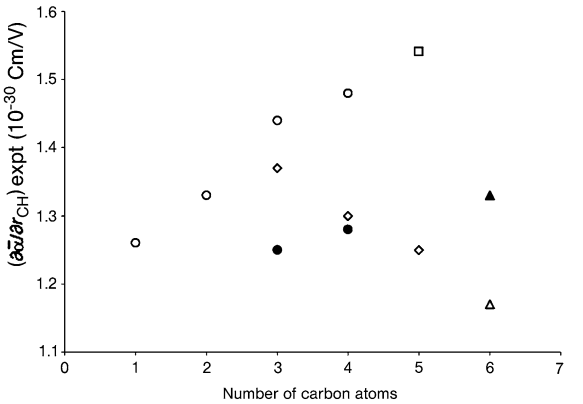


Fig. 4.2 Values of $\partial\bar{\alpha}/\partial r_{CH}$ from experimental Raman trace scattering intensities for methane, ethane, propane, *trans-n*-butane, bicyclo-[1.1.1]-pentane, and cyclohexane. H_{ip} (○), H_{op} (●), H_m (◇), H_{bh} (□), H_{ax} (△), H_{eq} (▲). Data from Refs. [6, 8–10, 24].

scattering is nearly invariant to isotopic substitution, enabling us to use the absolute intensity data for several isotopomers. Through force-field analysis, the derivatives are converted to internal coordinates $(\partial\bar{\alpha}/\partial r)$, where r identifies some unique CH or CC bond. These data properly report the change in the *molecular* polarizability associated with the displacement of an individual bond. Numerous experimental results demonstrate conclusively that $(\partial\bar{\alpha}/\partial r)$ for CH and CC stretches are strongly dependent on location and orientation, on molecular size, and on any strain or steric hindrance (Fig. 4.2).

4.2.1.2 Theoretical Modeling of Raman Scattering Intensities: What We Did and Why

The intriguing patterns discovered from experiment and supported by ab initio modeling [17–24] raised several new questions, most importantly, whether variation in $\partial\bar{\alpha}/\partial r_{CH}$ might increase with molecular size and complexity. We devised a survey set of thirty-five hydrocarbons comprising the all-*trans* conformers of *n*-alkanes up to C_{15}, *iso*-butane, six cycloalkanes and methylcycloalkanes, eight bicycloalkanes, four propellanes, and a tetracyclane [20]. These molecules represent several homologous series and also explore the full range of bonding in saturated hydrocarbons. The values of $\partial\bar{\alpha}/\partial r_{CH}$ for each symmetrically unique CH bond were obtained at the Hartree–Fock (HF) level of theory with the D95(d,p) basis set. Our combined experimental/theoretical approach may be summarized as:

- Use ab initio calculations to model known experimental data.
- Use QTAIM to analyze the wavefunctions, and to interpret and understand experiments.
- Perform new ab initio calculations to probe for and to predict patterns for untested molecules for which interesting behavior is expected.

- Use QTAIM to test our classifications, and to help us identify and understand what is transferable (guiding principles based on structure and composition) and what is not ($\partial\bar{\alpha}/\partial r$).

The highly strained bicyclo-[1.1.1]-pentane emerged as a likely example of a simple, highly symmetric molecule that would have dramatically different $\partial\bar{\alpha}/\partial r_{CH}$ for the CH bonds, with an absolute $\partial\bar{\alpha}/\partial r_{CH}$ for the bridgehead CH near the extreme of the values we have surveyed. That the experimental results perfectly matched the theoretically predicted values has been a very satisfying endorsement of our approach [24], enabling classification of the variations in $\partial\bar{\alpha}/\partial r_{CH}$ according to a few geometrical and chemical factors outlined in this chapter. We then completed extensive ab initio studies on other molecules to gain a larger perspective on the problem and to create a set of guidelines to describe what may be expected in other, untested situations, including some additional saturated hydrocarbons and some alkenes and silanes, for which we calculate the derivatives for C–H, Si–H, C–C, and Si–Si bonds.

It is, of course, of limited use for a theory to merely reproduce experiment if it does not offer reason or insight or predictive power in addition to replication. QTAIM analysis of the molecular wavefunction offers us the desired level of information from which insight may be gleaned, and from which new physical intuition may be derived. The complex multidimensional wavefunction yields to the familiar real-space world of the electron density. The simple representation of bonding by single, double and triple lines is replaced by the richness of a complete cataloguing of the electron distribution throughout a molecule.

The fundamental reasons for departure of a given property from an expected trend can often be difficult to identify. For molecules, we wish to scrutinize the various competing and complementary contributions to molecular properties. One of the triumphs of QTAIM has been to faithfully recover the functional groups of observational chemistry. Functional group transferability, however, is an outcome of QTAIM analysis for those instances when the functional group has experimentally-observed nominal transferability. Molecular structure consists of many elements, all of which are captured in the topology of the electronic charge distribution. We have many descriptors for these elements, for example strain and steric hindrance, the effects of which are local in nature yet affect the properties of the whole molecule. QTAIM tells us exactly how this occurs.

4.3 Methodology

Our purpose in Sections 4.1 and 4.2 was to make very clear that experiment has shown there are significantly different Raman scattering intensities from even such a seemingly robust/stolid/nondescript functional group as CH. We use QTAIM, which provides a real-space atomic-level description of the fundamental

molecular properties, to ascertain how the interconnectedness of the atoms and their relative structural arrangements lead to these dramatic changes in intensity from molecule to molecule. The main steps are summarized here; details, examples, advice and cautionary tales follow.

Step 1: Calculate α at the optimized geometry of the molecule in question. *Check*: compare α_{calc} with α_{expt} where possible

Step 2: Calculate α again for geometries in which the atoms have been displaced along a symmetric stretch coordinate,
e.g.: CH_4, let all $r_{CH} = r\,(\text{optimized}) \pm 0.01\,\text{Å}$

Step 3: Calculate $\Delta\alpha$ for this symmetry coordinate by finite difference and convert to $\Delta\alpha/\Delta r_{CH}$. *Check*: compare $(\Delta\alpha/\Delta r_{CH})_{calc}$ with $(\Delta\alpha/\Delta r_{CH})_{expt}$ where possible

Step 4: Perform QTAIM analysis on the wavefunctions obtained in Steps 1 and 2

Step 5: Recover α_{QTAIM} and $(\Delta\alpha/\Delta r_{CH})_{QTAIM}$ from atomic charges and atomic moments. *Check*: compare α_{QTAIM} and $(\Delta\alpha/\Delta r_{CH})_{QTAIM}$ with α_{calc} and $(\Delta\alpha/\Delta r_{CH})_{calc}$

Step 6: Examine the QTAIM atomic properties to gain insight into the patterns that emerge

4.3.1 Modeling α and $\partial\alpha/\partial r$

We have relied most heavily on the Gaussian [25] software but this is not a requirement if the wavefunction is accessible to the QTAIM software, typically AIM2000 [26]. Criteria for setting up a successful calculation and options for pushing for the highest accuracy with respect to experiment were recently reviewed [23]; the main points are presented here.

The basis set must be large and include diffuse and polarization functions to reproduce experimentally observed behavioral trends. In our earliest calculations we employed the D95(d,p) basis set because it met this criterion while still being small enough to enable calculations on molecules such as cyclohexane [18]. In the Gaussian implementation, α is calculated by use of the analytical coupled–perturbed Hartree–Fock (CPHF) method. More recently, we explored much larger basis sets using a coupled cluster method (CCSD(T)) and applied time-dependent density-functional theory to B3LYP hybrid density functionals [23, 24]. We obtained benchmark level calculations that accurately reproduce the Raman intensities for bicyclo-[1.1.1]-pentane to within experimental error.

The first criterion to be met is that the α determined computationally should be in good agreement with that found experimentally. To obtain calculated values that are within experimental error, we have found it adequate to use B3LYP with the aug-cc-pVTZ basis set. It is not necessary to demand such high absolute accu-

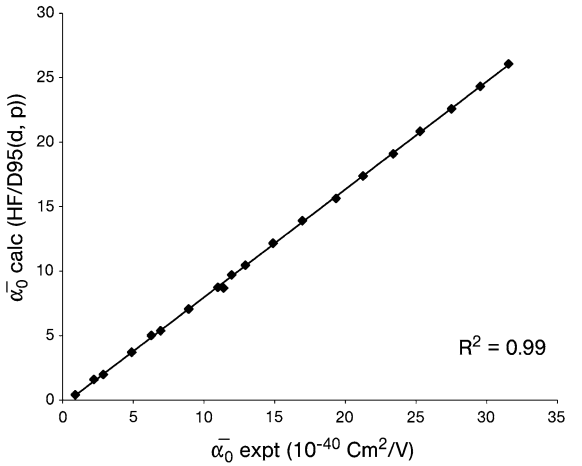


Fig. 4.3 Comparison with experimental values of mean molecular polarizability calculated at the HF/D95(d,p) level. Data from Refs. [22, 27].

racy to reproduce experimentally observed *trends*. Much of our early work was successful with the lower level HF/D95(d,p) calculations (Fig. 4.3) [21, 27]. For those readers who are interested in pushing the limits of accuracy, in Section 4.4 we present a detailed example of a high-level calculation that will match experimental data.

Having verified that α_{calc} is satisfactory, we approximate $\partial\bar{\alpha}/\partial r$ by finite difference:

$$\frac{\partial\bar{\alpha}}{\partial r} \approx \frac{\Delta\bar{\alpha}}{\Delta r} = \frac{1}{n}\left(\frac{\bar{\alpha}_{+\Delta r} - \bar{\alpha}_{-\Delta r}}{2\Delta r}\right) \tag{4}$$

where the change in bond length, Δr, is typically $\pm 1.00 \times 10^{-12}$ m; n = number of symmetrically equivalent bonds stretched or contracted, and $\bar{\alpha}_{\pm\Delta r}$ denotes the mean molecular polarizability at the stretched or contracted geometry. The accuracy of the central-difference derivative formula is second-order in step size; we have confirmed that good values are produced even for steps of $\Delta r = 1.00 \times 10^{-11}$ m [28]. The step size that we recommend here is a typically accepted value for this type of finite difference model.

4.3.2
Recouping α From the Wavefunction, With QTAIM

According to QTAIM, the dipole induced by the applied field is the result of charge transfer (CT) between atomic basins and changes in atomic dipole (AD)

within a basin. In practical terms, $\alpha_{i,j}(\Omega)$, the atomic contribution to the i, jth entry in the polarizability tensor, is expressed as:

$$\alpha_{i,j}(\Omega) = \frac{(N_i - N_0)r_j}{E_i} + \frac{\mu_i - \mu_0}{E_i} \tag{5}$$

where N_i and N_0 are the atomic electron populations when E is applied in the ith direction and when $E = 0$, respectively, for atom Ω located at r_j $(j = x, y, z)$. Atomic population is the atomic number minus the charge. The AD term is found from the change in the atomic first moment, μ, with and without the applied field. The $\alpha_{i,j}(\Omega)$ are summed over all atoms to give the molecular tensor α.

4.3.3
Recovering $\partial\alpha/\partial r$ From QTAIM

In other chapters we have seen the incredible insight into structure and its underlying basis, bonding, gained by QTAIM analyses. In essence, this approach puts skin on the bones of the molecular skeleton; we want, however, to animate our understanding of bonding, to go beyond the question of static equilibrium structure and delve into the response of the molecule to perturbation, as seen in the changes in the topology of the electron density. Short of ionization, we know that the electrons will remain in the molecule. This is rather unsatisfying. We can look a little more deeply and report that, in the presence of a perturbing laser field, the electrons would redistribute themselves within the molecule. As in the static example, an atomic-level description is convenient and powerful, and provides a rigorous connection between theory and experiment. The picture now becomes one of electronic flow throughout a molecule and of reorientation of charge within atoms.

The recovery of the derivative from the wavefunctions that we have created is a straightforward process, based on Eqs (4) and (5). For each atom, we can examine:

$$\frac{\partial\alpha_{i,j}}{\partial r} \cong \frac{\Delta\alpha_{i,j}(\Omega)}{\Delta r} = \frac{(N_i - N_0)r_j}{\Delta r \times E_i} + \frac{\mu_i - \mu_0}{\Delta r \times E_i} \tag{6}$$

The right hand side of the equation quantifies the changes in the CT and AD terms at each atom ($\Delta CT(\Omega)/\Delta r$ and $\Delta AD(\Omega)/\Delta r$), in response to the molecular vibration and the applied field. We will discuss the patterns and trends that QTAIM reveals in Sections 4.5 (α) and 4.6 ($\partial\bar{\alpha}/\partial r$).

4.4
Specific Examples of the Use of AIM2000 Software to Analyze Raman Intensities

We next carry out the step-by-step procedure of Raman intensity analysis through QTAIM for H_2 and CH_4. These molecules are small enough to enable excellent

calculations relatively quickly while still demonstrating the major outcomes. It is important to distinguish between the accuracy or method of the calculation and the actual convergence level of that chosen method. When α is computed numerically (through energy or dipole derivatives) using small applied fields, the results are sensitive to the SCF convergence, especially when the derived polarizabilities are computed from small displacements in bond lengths, chosen to better approximate $\partial\bar{\alpha}/\partial r$. On the basis of error analysis, we recommend a computationally inexpensive SCF convergence of 10^{-11} rather than 10^{-8}, a typical default value.

4.4.1 Modeling α in H_2

For H_2 with only two electrons and one geometric parameter, it is possible to perform a very high-level calculation on an ordinary PC; CCSD for a two-electron system is equivalent to the full configuration interaction method. With that, we use the aug-cc-pVQZ basis set, a large correlation consistent basis (quadruple zeta) with diffuse and polarization functions. The CCSD value of $\bar{\alpha} = 0.856 \times 10^{-40}$ $Cm^2\ V^{-1}$ compares very well with experiment (0.858×10^{-40} $Cm^2\ V^{-1}$) [29], as per our check of Step 1. Following Steps 2 and 3, we obtain a CCSD value for $\Delta\bar{\alpha}/\Delta r$ of 1.36×10^{-30} Cm V^{-1}, again in excellent agreement with several experimental values (1.36×10^{-30} Cm V^{-1}) [30].

For the QTAIM calculations we use $E = 0.001$ au, the same field strength as in the numerical CCSD Gaussian calculation, to facilitate direct comparisons. It is small enough to minimize the effect of higher-order polarizabilities, and tends to give excellent agreement with analytical (CPHF) calculations where the method allows. One must usually compute wavefunctions for $E_0 = 0$ and with fields E_x, E_y, and E_z, for each geometry, and average the results obtained from fields applied in both directions along each axis. For H_2 we had aligned the principal axis (bond axis) with the z-axis, thereby diagonalizing the polarizability tensor and making $\alpha_{xx} = \alpha_{yy}$. Thus for each unique geometry ($\Delta r = 0, \pm 1.00 \times 10^{-12}$ m), we need only to compute a reference wavefunction file with E_0, E_x, and E_z applied fields (the + and − directions being equivalent).

It must be stressed that analysis of Raman intensities requires the maximum accuracy from QTAIM integrations, because we are looking for the effects of very small perturbations (electrostatic and geometric) on the atomic electronic properties. A relatively small error of 10^{-3} au in $\mu(\Omega)$ becomes an error of 1 au in $\alpha(\Omega)$, in turn producing huge errors in the calculated $\partial\bar{\alpha}/\partial r$; any insights gained would be suspect. With AIM2000 we have found the greatest accuracy to be obtained by integrating in "natural coordinates". We set the relative and absolute integration accuracy values to 1×10^{-6}, which are two orders of magnitude tighter than the default and is at the limit of the numerical stability of the current implementation. The beta-sphere radius is set equal to the distance from the nuclear attractor (NA) (not necessarily exactly at the nuclear position) to the nearest bond critical point. Finally we triple the default maximum path length "s" to

Table 4.1 QTAIM integration results for H_2 at equilibrium geometry (in atomic units).

	Ω	Atomic charge	μ_x	μ_y	μ_z	Atomic energy
E_0	H1	0.0002989	−0.0000047	−0.0001498	0.0945175	−0.5666
	H2	0.0003008	0.0000004	0.0001436	−0.0945188	−0.5666
E_x	H1	0.0002061	0.0023122	−0.0001386	0.0944551	−0.5667
	H2	0.0002049	0.0023128	0.0001412	−0.0944540	−0.5667
E_z	H1	−0.0041055	0.0000028	−0.0001437	0.0946650	−0.5683
	H2	0.0046799	−0.0000005	0.0001420	−0.0943614	−0.5650

3×10^6. The path length has never been a limiting factor for the molecules considered. With these values it is common to recover α to better than 99%; this level of completeness enables, at the very least, semi-quantitative recovery of Raman intensities.

In Table 4.1 we show the integration data for H_2 at equilibrium geometry for E_0, E_x, and E_z (note that all data are in au unless otherwise stated). For E_0, the charge on each atom is zero to within a few parts in 10^4; the x and y atomic dipoles cancel to within a few parts in 10^6. Although the z-component of the *molecular* dipole should be zero, each *atom* has a non-zero dipole reflecting the shape of the densities around the nuclei; these are equal and opposite. The net magnitude of the molecular dipole is less than 1×10^{-5}. This can be regarded as a high-quality integration, because the dipole error would only affect calculated polarizabilities by 1×10^{-2}. For E_x, symmetry excludes the possibility of charge transfer; only an induced dipole in the x-direction is possible. The x-coordinate of the density maxima (NA) is sufficiently distorted by the field to move slightly away from the nuclear position, $r_x(\mathrm{NA}) = 3.586 \times 10^{-5}$, $r_x(\mathrm{H}) = 0$. This is typical of light atoms. One must be careful to use the position of the NA and not the nucleus when applying the formulas for atomic electronic properties. For $E_z = 0.001$ au, there is both CT between hydrogen atoms and a nonzero AD (distortion of the density) in each atomic basin. Average atomic contributions to the α_{xx} and α_{zz} elements, in au, are computed from Eq. (5): $\mathrm{CT}(\Omega)_{xx} = 0$, $\mathrm{AD}(\Omega)_{xx} = 2.312$; $\mathrm{CT}(\Omega)_{zz} = 2.996$, $\mathrm{AD}(\Omega)_{zz} = 0.152$; after summation and conversion we get $\bar{\alpha}(H_2) = 0.854 \times 10^{-40}$ $\mathrm{Cm^2\ V^{-1}}$. The atomic average is equivalent to averaging the field in both directions along the axis; we find an excellent QTAIM recovery of 99.7601%.

A word of caution is necessary here. The robustness of electronic structure codes is usually much greater than that of available QTAIM integration software (for the accuracies we require). This means that an electronic structure calculation should always converge to precisely the same energy (and properties) for symmetry-equivalent species. In contrast, the nature of the QTAIM integration

grid *could* produce orientation-dependant results. With our very accurate values we could not find any instances of this, despite using much stronger fields to distort the charge densities, and intentionally integrating in favorable and unfavorable orientations, data not shown.

4.4.1.1 Modeling $\Delta\bar{\alpha}/\Delta r$ in H_2

To complete the analysis of Raman intensities for H_2, one must repeat the steps above for stretched and contracted geometries. Our usual approach is to displace the nuclei by 0.01 Å; the error in α_{QTAIM} (less than 2%) is, however, still enough to cause problems with interpretation of $\partial\bar{\alpha}/\partial r$ (QTAIM recovery error of >30%). We repeated the calculation for $\Delta r = \pm 0.04$ Å and obtained much better agreement between $\pm\Delta r$, and better overall recovery. The average atomic contributions to α are shown for $\Delta r = +0.04$ Å, to highlight the orientation dependence of the changes, $i = x, y, z$:

$$
\begin{aligned}
&x \text{ field}\\
&\Delta CT(H)_{ix} = (0.000 \quad 0.000 \quad 0.002)\\
&\Delta AD(H)_{ix} = (0.132 \quad -0.002 \quad -0.002)\\
&z \text{ field}\\
&\Delta CT(H)_{iz} = (0.000 \quad 0.000 \quad 0.205)\\
&\Delta AD(H)_{iz} = (-0.003 \quad -0.004 \quad 0.078)
\end{aligned}
\tag{7}
$$

For H_2, the perpendicular contribution (α_{xx}) is necessarily entirely AD; CT is the dominant contribution to α along the bond (z axis) and the largest tensor component overall. The recovered value $\Delta\bar{\alpha}/\Delta r_{QTAIM}$ for H_2 (1.48×10^{-30} Cm V^{-1}) is in reasonable agreement with that calculated in Gaussian (1.36×10^{-30} Cm V^{-1}), which is identical with the experimental value [30].

4.4.2 Modeling α and $\Delta\bar{\alpha}/\Delta r$ in CH_4

For this larger system with ten electrons, we have chosen the B3LYP hybrid density functional method with the aug-cc-pVTZ basis set. This is very computationally affordable for medium and even larger systems. The computed $\partial\bar{\alpha}/\partial r_{CH}$ is in perfect agreement with experiment [31] (Steps 1–3, and below). Another advantage is that analytical polarizabilities are available by use of the CPHF method. This enables us to compare numerical polarizability results obtained from energy or dipole derivatives with the analytical values. With its Td symmetry, methane has only one unique, nonzero polarizability tensor element. We chose to orient the molecule so that the hydrogen atoms appear at the corners of a cube. The QTAIM results for the equilibrium geometry are shown in Table 4.2. The x, y, and z directions are equivalent; our electronic structure calculations produce the same results for all field directions (and are unchanged for a C_{3v} geometry with

Table 4.2 QTAIM integration results for methane at equilibrium geometry for E_0 and E_x (in atomic units).

	Ω	Atomic charge	μ_x	μ_y	μ_z	Atomic energy
E_0	C	0.0219197	−0.0004455	−0.0000028	0.0000110	−38.0411
	H1	−0.0053341	0.0854614	0.0853861	0.0853876	−0.6243
	H2	−0.0049950	−0.0860062	−0.0855185	0.0854416	−0.6241
	H3	−0.0050035	−0.0860009	0.0855167	−0.0854399	−0.6241
	H4	−0.0053182	0.0854648	−0.0853973	−0.0853895	−0.6243
E_x	C	0.0219291	−0.0004116	−0.0000410	−0.0015141	−38.0411
	H1	−0.0079645	0.0853935	0.0853224	0.0869203	−0.6253
	H2	−0.0076480	−0.0859373	−0.0854687	0.0869380	−0.6251
	H3	−0.0023261	−0.0860613	0.0856022	−0.0839633	−0.6230
	H4	−0.0026376	0.0855481	−0.0854562	−0.0839168	−0.6232

one CH pointed along the z axis, not shown). Thus, we need only perform two QTAIM analyses (E_0 and E_z) at each of $\Delta r = 0, \pm 0.01 \times 10^{-12}$ m (Table 4.3). The individual CT_{xz} terms are large but cancel on summation, owing to symmetry; individual AD_{xz} terms are quite small, and also cancel. For the zz terms, we see the interesting effect of CT from one end of the molecule to the other, with no net CT to the carbon atom. Carbon, however, has an AD opposing this effect, and the hydrogen atoms a large AD terms, reflecting the nonspherical density around each H, in the field direction.

The change in atomic contributions to α at the equilibrium, stretched, and contracted geometries are illustrative of the insights into Raman intensities and charge flow in general that are possible through QTAIM analysis (Table 4.4).

Table 4.3 QTAIM analysis of CT(Ω) and AD(Ω) to $\alpha_{iz}(CH_4)$ at E_z, $\Delta r = 0$ (in atomic units: $a_0{}^3$).

Ω	$CT_{xz}(\Omega)$	$CT_{yz}(\Omega)$	$CT_{zz}(\Omega)$	$AD_{xz}(\Omega)$	$AD_{yz}(\Omega)$	$AD_{zz}(\Omega)$	α_{xz}	α_{yz}	α_{zz}
C	0.0	0.0	0.0	0.03386	−0.03819	−1.52518	0.03386	−0.03819	−1.52518
H1	3.07854	3.07854	3.07864	−0.06794	−0.06371	1.53278	3.01060	3.01483	4.61142
H2	−3.10500	−3.10509	3.10518	0.06883	0.04983	1.49646	−3.03617	−3.05526	4.60164
H3	3.13346	−3.13355	3.13345	−0.06038	0.08544	1.47667	3.07308	−3.04811	4.61012
H4	−3.13723	3.13723	3.13713	0.08329	−0.05882	1.47276	−3.05395	3.07842	4.60989
Σ_Ω	−0.03023	−0.02287	12.45440	0.05766	−0.02545	4.45349	0.02743	−0.04832	16.90789
B3LYP/aug-cc-pVTZ							0.0	0.0	17.02210
Recovery by QTAIM %							100.027	99.952	99.329

Table 4.4 QTAIM recovery of α and $\Delta\alpha/\Delta r$ for CH_4, where $\partial\bar{\alpha}/\partial r \equiv \partial\alpha_{zz}/\partial r$.

Parameter(units)[a]	Δr(Å)	B3LYP	QTAIM	Recovery (%)	Expt.
α_{zz} $(a_0{}^3)$	−0.01	16.7200	16.5572	99.026	
α_{zz} $(a_0{}^3)$		17.0221	16.9079	99.329	
α_{zz} $(a_0{}^3)$	+0.01	17.3288	17.1638	99.048	
$\bar{\alpha}$ $(Cm^2\ V^{-1})$	0	2.807×10^{-40}	2.787×10^{-40}		2.85[b]–2.93[c] $\times 10^{-40}$
$\Delta\alpha/\Delta r_-$ $(Cm\ V^{-1})$		1.245×10^{-30}	1.055×10^{-30}		
$\Delta\alpha/\Delta r_+$ $(Cm\ V^{-1})$		1.264×10^{-30}	1.446×10^{-30}		
$\Delta\alpha/\Delta r$ $(Cm\ V^{-1})$	average	1.255×10^{-30}	1.250×10^{-30}	99.64	1.26×10^{-30}[d]

[a] Conversion factor $1.648777631 \times 10^{-41} * (a_0)^3 = Cm^2\ V^{-1}$.
[b] Ref. [27].
[c] Ref. [33].
[d] Ref. [32].

The average atomic contributions for the E_z field at $\Delta r = 0$ are:

$$\begin{aligned} CT_{iz}(C) &= (0 \quad 0 \quad 0) \\ CT_{iz}(H) &= (-0.008 \quad -0.006 \quad +3.114) \\ AD_{iz}(C) &= (+0.034 \quad -0.038 \quad -1.525) \\ AD_{iz}(H) &= (-0.002 \quad -0.003 \quad +4.608) \end{aligned} \tag{8}$$

A significant contribution to $\Delta\bar{\alpha}/\Delta r_{CH}$ is the decrease in the opposing AD in the carbon atom by 0.045 au along with increases in CT(H) and AD(H). This type of analysis for systems with unusual charge flow (reflected in the Raman intensities) will reveal the structural factors that govern the most extreme behavior.

4.4.3
Additional Exercises for the Interested Reader

Given the limitations in QTAIM integration accuracy, it is important to establish the best geometric step size for studying $(\partial\bar{\alpha}/\partial\mathbf{r})$. Using the data provided below, compute $(\Delta\bar{\alpha}/\Delta\mathbf{r})$ for methane and H_2 using the first-order accurate forward–difference method for:

(a) stretched and equilibrium geometry, and
(b) equilibrium and contracted geometry.

Answers:

(a) H_2: 1.362×10^{-30} Cm V^{-1} methane: 1.264×10^{-30} Cm V^{-1}; and
(b) H_2: 1.352×10^{-30} Cm V^{-1} methane: 1.245×10^{-30} Cm V^{-1}

Compare these results with those from the second-order accurate central-difference method using the stretched and contracted data.
Answers: H_2 $\Delta\bar{\alpha}/\Delta\mathbf{r}$: 1.357×10^{-30} Cm V^{-1} methane: 1.255×10^{-30} Cm V^{-1}

Compute the second derivatives using the three-point second-derivative formula and the data below.
Answers: H_2: 9.735×10^{-21} C V^{-1} methane: 1.895×10^{-21} C V^{-1}.

Run electronic structure calculations computing polarizabilities for methane and H_2 at larger bond displacements (e.g. ± 0.02 Å, ± 0.04 Å, ... ± 0.1 Å).

How do the values obtained for $(\Delta\bar{\alpha}/\Delta\mathbf{r})$ compare with those from the smaller step size? What is the largest step-size that preserves $(\Delta\bar{\alpha}/\Delta\mathbf{r})$ to within 2%? Could the largest accurate step-size be estimated from the small step-size second-derivative result?

Data: $\bar{\alpha}$ (au) at Contracted/Equilibrium/Stretched geometries (± 0.01 Å)
H_2: 5.113/5.195/5.277.
Methane: 16.720/17.022/17.329.

4.5
Patterns in α That Are Discovered Through QTAIM

Molecular polarizability to describes the displacement of the electrons in a molecule when it is placed in an electric field. In the language of QTAIM, this displacement has two parts: transfer of charge from one atomic basin to another and polarization of charge within an atomic basin. The challenge for us is that the properties we are examining are extraordinarily sensitive to our (in)ability to numerically recover the full information available in the wavefunction. It is important to emphasize that this is a numerical and implementation problem rather than a fundamental problem with the QTAIM theory. Fortunately, we do not have to push the calculation levels to the limit in Section 4.4 to uncover these patterns. The results are worthwhile, as is seen in the QTAIM analysis of the polarizabilities of the simple alkanes, for example methane, ethane, and propane [17]. Initially, we only used the D95(d,p) basis at the HF level, yet we discovered the patterns of behavior [17–21] identical with those revealed in Section 4.4 and calculated with larger basis sets [27]. In ethane and propane we see the emergence of the patterns in α that underlie those in $\partial\bar{\alpha}/\partial$r [17].

In Fig. 4.4, we see the effect of E_z on propane. Just as in CH_4, the CT term is greatest along the field axis and is essentially a transfer between the atoms at the extreme ends of the molecule. In propane, these are the H atoms that lie in the skeletal plane, denoted H_{ip} (ip = in plane). There is a smaller though still significant transfer between the methyl hydrogen atoms that lie out of the plane, denoted H_{op} (op = out of plane) and between the terminal methyl carbons. The methylene group (CH_m) does not participate. As might be expected, integration over the atomic basins of each atom shows that the charge is transferred symmetrically, creating a CT dipole that opposes the applied field. The story changes when we examine the atomic dipoles. The AD(H) are aligned with those created

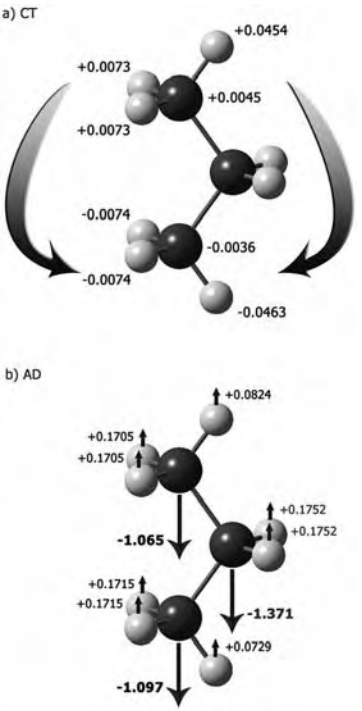


Fig. 4.4 QTAIM analysis of response for an electric field applied to propane in the *z* direction. (a) Charge transfer (CT) and (b) atomic dipole (AD) contributions (10^{-40} Cm^2 V^{-1}) to the α_{zz} term in the molecular polarizability of propane. By convention, the *z*-axis is defined as the axis with maximum polarizability. Data were calculated at the HF/D95(d,p) level, from Ref. [17].

through CT. However, the ADs of all three Cs that form the inner skeleton of the molecule are directed in the opposite sense. The molecular polarizability is the net result of these opposing contributions.

We found the identical pattern in each field direction and in every molecule we have analyzed with QTAIM. The amount of charge transferred between terminal H atoms increases with chain length, while the opposing ADs in the carbon skeleton also increase [19]. In cyclohexane [18], when the field is applied along the nominal ring plane, the CT(H_{eq}) terms are very large. For the field perpendicular to the ring plane, the CT between the H_{ax} is significant; the CT contribution in Eqs. (5) and (6) is still smaller, however, because the distance over which the charge is transferred is smaller. The picture that emerges is similar to the surface charge polarization of an idealized dielectric body of uniform positive and negative charges. The surface polarization overrides the effect of the external field on the interior. The CT on the outer H atoms is countered by an AD formed by reorganization of charge within the carbon atomic basins.

4.6
Patterns in $\partial\alpha/\partial r_{CH}$ That Apply Across Different Structures, Conformations, and Molecular Types: What is Transferable?

In this section, the emphasis is focused on the dynamic nature of $\partial\alpha/\partial r$ compared with the more static nature of the polarizability. In exploring $\partial\alpha/\partial r$, we seek local details rather than merely global information and QTAIM provides the means with which to examine the molecular wavefunction. The derivatives are not transferable, so what is? The answer lies in the patterns revealed by QTAIM analysis.

Experimental data on absolute Raman trace scattering intensities are the acid test for the validity of the theoretical models; such experiments are tractable, however, only for small, highly symmetrical molecules that can be studied easily in the gas phase. Experimental $\partial\bar{\alpha}/\partial r$ have been obtained for methane, ethane, propane, butane, cyclohexane, ethene, ethyne, and, most recently, for bicyclo-[1.1.1]-pentane [6, 8–13]. Relative values for $\partial\bar{\alpha}/\partial r_{CH}$ have been determined for *n*-pentane, which occurs as a mixture of *trans–trans*, *gauche–trans* and *gauche–gauche* conformers in the gas phase at room temperature [19]. Our detailed QTAIM analyses have been directed to these molecules. We have also completed survey studies, Steps 1–3 above, on over 50 molecules, to identify trends in calculated $\Delta\bar{\alpha}/\Delta r$ values and to find appropriate candidates for further detailed experimental and theoretical analysis [20–22].

4.6.1
Patterns in $\Delta\bar{\alpha}/\Delta r_{CH}$ Revealed by QTAIM

From our QTAIM analysis, we have seen that alkanes behave like tiny pieces of polarizable material, with end-to-end transfer of charge opposed by internal dipoles. Although this behavior can be trivially understood in a gross sense, what is of greater interest is to understand the effect of the molecular framework on the flow of charge and the rearrangement of atomic electron densities. We are studying the behavior of a charge cloud bound by nuclear attractors and distributed across distinct atomic basins. With the data in hand we can now explore the dynamics of the electron density during a molecular vibration.

4.6.1.1 QTAIM Analysis of $\Delta\bar{\alpha}/\Delta r_{CH}$ in Small Alkanes

Complete QTAIM analysis has been conducted for methane, ethane, propane, the *trans*, *trans* conformation of *n*-pentane, and cyclohexane [17–19]. The first thing we note is that the change in $\bar{\alpha}$, the mean *molecular* polarizability associated with the stretch of any CH bond, arises from changes distributed throughout the molecule [17–19]. The complexity of the response is illustrated for the three different CH stretch modes in propane (Fig. 4.5) [17]. The numbers beside each atom are the atomic contributions (Eq. 6) divided by the number of equivalent bonds, to give $\Delta\bar{\alpha}(\Omega)/\Delta r_{CH}$, i.e. the derivative per CH bond (all $\Delta\bar{\alpha}/\Delta r$ are reported in units of 10^{-30} Cm V^{-1}). The contributions are largest for the atoms of the bond being stretched, but the contributions from the remainder of the molecule cannot be

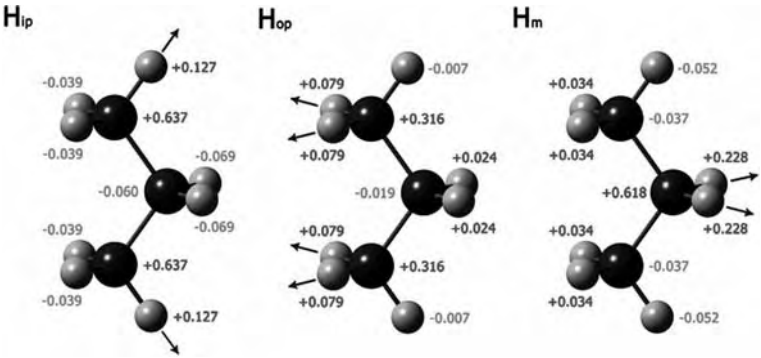


Fig. 4.5 QTAIM analysis of the atomic contributions to $\Delta\bar{\alpha}/\Delta r_{CH}$ in propane for (a) methyl in-plane CH stretch, (b) methyl out-of-plane CH stretch, and (c) methylene CH stretch. Data were calculated at the HF/D95(d,p) level, from Ref. [17].

ignored. For the stretch of CH_{ip}, there is a net increase in the atomic contributions to $\bar{\alpha}$ from the atoms of the bond being stretched, but a decrease for the remaining atoms. For the CH_{op} and CH_m bond stretches, the large positive atomic contributions to $\Delta\bar{\alpha}/\Delta r_{CH}$ are again from the atoms of the stretched bond, but the H_{op} and H_m terms are positive for both.

When the atomic contributions are broken down into changes in the CT and AD terms [17], the results are intriguing, although no overriding pattern can be discerned from this small set. For all three bond stretches, the AD term is larger than the CT.

- For the CH_{ip} stretch, the $\Delta AD/\Delta r_{CH}$ term is positive for both H_{ip} and C_{ter} (terminal carbon, to which it is bonded) but negative for other atoms. The $\Delta CT/\Delta r_{CH}$ term increases for C_{ter} only.
- For the CH_{op} stretch, the $\Delta AD/\Delta r_{CH}$ term is large and positive for H_{op} and C_{ter} but small and negative for the other atoms. The $\Delta CT/\Delta r_{CH}$ term is almost zero for all but the C_{ter}, for which it is only one quarter the magnitude of the value for the CH_{ip} stretch.
- For the CH_m stretch, $\Delta AD/\Delta r_{CH}$ terms are large and positive for H_{op} and C_m (methylene carbon) and negative for the other atoms. The $\Delta CT/\Delta r_{CH}$ term is again quite small for most atoms. It is larger and positive for C_m and, surprisingly, for H_{op}.

For methane, ethane and propane, the $\Delta AD/\Delta r_{CH}$ contributions to $\Delta\bar{\alpha}/\Delta r_{CH}$ are quite similar [17]. As we go from methane to ethane to the CH_{ip} stretch in propane, however, the $\Delta CT/\Delta r_{CH}$ of the bonded carbon increases steadily, from 0 in methane, to 0.286 in ethane, to 0.887 in propane. The same term is only

0.272 and 0.354 for the bonded carbon in the CH_{op} and CH_m stretches, respectively. For the CH_{ip} stretch in propane, the *xx* and *yy* terms are negligible; almost the entire contribution comes from the *zz* term: 0, −0.291, and 2.952, respectively. This is intuitively reasonable, because the CH_{ip} stretch is almost entirely in the *z* direction, thus one expects the change in the polarizability to be greatest in this direction.

The trend continues when we compare methane, ethane, and propane with the larger molecules cyclohexane [18] and *n*-pentane [19]. We begin with the C_{ter} *zz* contribution to the CH_{ip} stretch in pentane. At 1.12, the $\Delta AD/\Delta r_{CH}$ term is larger than for any of the smaller molecules; the most dramatic change, however, is found for the $\Delta CT/\Delta r_{CH}$ term, which is now 10.04. The corresponding terms (*xx* and *yy*) for the CH_{op} and CH_m stretches are an order of magnitude smaller.

The disparate intensities from CH_{ax} and CH_{eq} in cyclohexane set our studies in motion, so it is very satisfying that QTAIM analysis has provided the key that resolves the puzzle. In our arrangement, the CH_{eq} lie in the *xy* plane and, for the CH_{eq} stretch, the $xx + yy$ contributions from the C atoms to $\Delta CT/\Delta r_{CH}$ combine to a total of 3.346. This is significantly greater than the corresponding contributions to $\Delta CT/\Delta r_{CH}$ for the CH_{ax} stretch. As in propane, this CT term is partly offset by negative contributions at the H_{eq}. The H_{eq} effectively lie at the end of a pair of short carbon ''chains'', like the H_{ip}. Although CH_{ax} is almost completely aligned with the *z*-axis, the carbon *zz* contribution is only 0.420. Despite differences in the detail, they are more similar to the CH_m in propane and pentane.

4.6.1.2 What Did We Learn From QTAIM That Can be Transferred to the Other Molecules?

In parsing the quantity of detail in the preceding sections, we have concluded that it is at least as important to consider the distance across which charge is transferred as to consider the absolute amount of charge transferred. For any CH bond stretch, the change in the molecular polarizability is greatest for the external field that is most closely aligned with the bond. The CH_{eq} in cyclohexane are structurally similar to the CH_{ip} bonds in propane and pentane, in that all are aligned with their carbon chains and effectively lie in the plane of the carbon atoms. Interestingly, these CH bonds are the shortest in the molecule and always produce the greatest Raman scattering intensity. The CH_{ax} in cyclohexane have more in common with the CH_{op} and CH_m bonds; they are in a plane orthogonal to the carbon framework and are typically slightly longer, with lower isolated CH stretching frequencies and weaker Raman scattering intensities (Fig. 4.1).

From experiment we know that $\partial\bar{\alpha}/\partial r\ (CH_{ip})$ will be larger than that for any other CH bond in a given molecule, and will increase with chain length. From QTAIM analysis we see that the distinguishing feature is the increase in the ΔCT contribution of the carbon atom. The molecular polarizability is greatest down the length of the chain; it is also most sensitive to alterations in this dimension. The ΔAD contributions are mainly confined to the atoms of the bond being

stretched. While they too increase with molecular size they do not exhibit, the dramatic differences found for the ΔCT(C).

These results form the first of our guiding principles in the Raman scattering intensities – the magnitude of $\partial\bar{\alpha}/\partial r_{CH}$ is highly dependent on bond location within the molecule, orientation relative to the carbon skeleton, and the spatial extent of that skeleton.

4.7 What Can We Deduce From Simple Inspection of $\partial\bar{\alpha}/\partial r_{CH}$ and $\partial\bar{\alpha}/\partial r_{CC}$ From Gaussian?

The larger survey calculations of $\partial\bar{\alpha}/\partial r$ using standard electronic structure calculations include normal and branched alkanes, a diverse selection of cycloalkanes and bicycloalkanes, hedranes and propellanes, and a few alkenes, alkynes, silanes, and extended sheets [22–24]. Some of these are shown in Fig. 4.6. One of our goals was to explore the potential variety of $\partial\bar{\alpha}/\partial r$ values that might exist, seeking extensions to the patterns found above and possible new patterns or factors to consider. A second goal was to identify candidate molecules for which extreme differences in $\partial\bar{\alpha}/\partial r$ might be observed and which might be amenable to experimental and QTAIM analysis. The CH and CC stretch in all-*trans* straight-chain alkanes up to pentadecane were modeled at the HF/D95(d,p) level [19–21] as well as the $\Delta\bar{\alpha}/\Delta r_{CH}$ for the *gauche*-butane, and the *gauche–trans* and *gauche–gauche*-pentane conformers. The $\Delta\bar{\alpha}/\Delta r_{CC}$ were calculated up to *n*-pentacosane ($C_{25}H_{52}$). We found that the variety of differences could be categorized into a few simple patterns that have been discussed thoroughly in our papers. Here we present specific examples to illustrate the factors that were identified.

4.7.1 Variations in $\partial\bar{\alpha}/\partial r_{CH}$ Among the Alkanes

The patterns discovered in the experimental data (Fig. 4.2) are a small part of the larger picture [19, 22, 23].

- The stretch of the CH_{ip} bond always produces the greatest change in $\bar{\alpha}$.
- The CH_{op} and CH_m bonds, orthogonal to the chain, are all quite similar, though there is a slow, regular decrease from the $\Delta\bar{\alpha}/\Delta r_{CH}$ in propane with increasing chain length.
- Where conformational change rotates the CH bonds to point toward each other, and toward the interior of the molecular skeleton, the value of $\Delta\bar{\alpha}/\Delta r_{CH}$ decreases. For CH bonds rotated into the plane of the carbon chain $\Delta\bar{\alpha}/\Delta r_{CH}$ increases (Fig. 4.7).

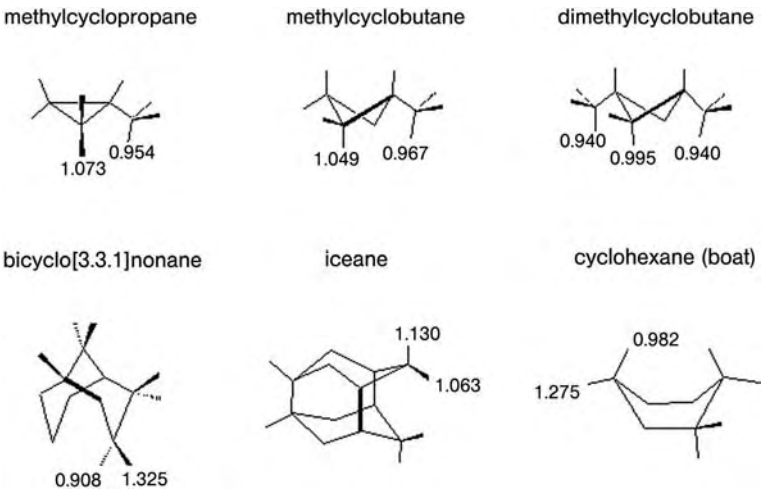
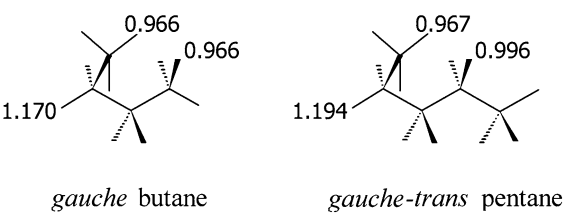


Fig. 4.6 Effect of conformation and structure on the magnitude of $\partial\bar{\alpha}/\partial r_{CH}$. Exo/endo differences are illustrated in the boat form of cyclohexane, steric hindrance in methylcyclopropane, methylcyclobutane, and 1,3-dimethylcyclobutane, and the combined effects of axial/equatorial and exo/endo are evident in some of the CH bonds of bicyclo-[3.3.1]-nonane and tetracyclo-[5.3.3.$1^{2,6}$.$0^{4,9}$]-dodecane, also called iceane. Data were calculated at the HF/D95(d,p) level, from Refs. [20, 21].

- At 1.118, $\Delta\bar{\alpha}/\Delta r_{CH}$ for the methine CH bond is largest in *iso*-butane. It has the "in-plane" alignment; the methyl CH bonds are in-plane (1.107) or out-of-plane (1.000) (Figure not shown).

QTAIM analyses have not been performed on most of these molecules. We speculate that the pattern of end-to-end CT, opposed by increased ADs at the carbons atoms will be the overriding factor. There is a gradual leveling with chain



gauche butane *gauche-trans* pentane

Fig. 4.7 Effect of molecular conformation on the magnitude of $\Delta\bar{\alpha}/\Delta r_{CH}$ in *gauche n*-butane and *gauche-trans n*-pentane. The steric hindrance imposed by internal rotation reduces the derivative from approximately 1.05×10^{-30} Cm V^{-1} for the methyl out-of-plane and methylene CH. Rotation of a CH into an end-of-chain in-plane position raises the derivative to a value close to those of the H_{ip} in all-*trans* butane and pentane (1.207 and 1.26, respectively). Data were calculated at the HF/D95(d,p) level, from Refs. [19, 20].

length, as a result of either the damping effect of intervening methylenes or the unfavorable nature of large CT over greater distances.

4.7.2
$\Delta\bar{\alpha}/\Delta r_{CH}$ in Cycloalkanes, Bicycloalkanes, and Hedranes

Alkane rings contain additional categories of CH bond, simplified as axial, equatorial, exo, endo and bridgehead. The CH_{ax} and CH_{eq} have already been addressed in cyclohexane. The points below represent the cycloalkane variations on the trends discovered in the straight chains (Fig. 4.6).

- The difference between CH_{ax} and CH_{eq} increases with ring size.
- Steric crowding and rotation such that bonds are directed toward the interior of the skeleton results in a decrease in $\Delta\bar{\alpha}/\Delta r$, whether the bonds are part of the ring or on methyl substituents.
- Exo/endo differences appear, as seen in the boat form of cyclohexane.
- Individual effects identified in simple structures combine in more complicated molecules.

Bicyclo-[3.3.1]-nonane and tetracyclo-[$5.3.1.1^{2,6}.0^{4,9}$]-dodecane (iceane) serve to illustrate the combined effects of axial, equatorial, exo, and endo characteristics. The most dramatic difference in $\partial\bar{\alpha}/\partial r$ was predicted and experimentally verified in bicyclo-[1.1.1]-pentane [24]. The bridgehead CH experiences the combined enhancement of the terminal position and the strain at the carbon atom, to give the largest $\partial\bar{\alpha}/\partial r_{CH}$ of any molecule in this category, irrespective of size.

4.7.3
Patterns That Emerge in $\Delta\bar{\alpha}/\Delta r_{CC}$ of Alkanes

The unique CC bonds in straight-chain alkanes and in simple cycloalkanes may be easily examined with HF and DFT methods [21–23]. On stretching the skeleton CC bonds we observe complete reversal of the patterns seen in the CH bonds. The magnitude of $\Delta\bar{\alpha}/\Delta r_{CC}$ is smallest when bonds at the outer ends of the chain are stretched, becoming larger as we progress towards the center of the chain. Even more curiously, the smallest derivative occurs at the second CC bond from the end, as if the terminal position imparts a slightly greater looseness to the charge density. This is not an artifact of the lower level basis set, because the calculations were repeated with B3LYP/aug-cc-pVDZ and identical patterns appeared (Fig. 4.8). With QTAIM analysis, we hope to gain insight that will explain these predictions.

For the ring and cage molecules $\Delta\bar{\alpha}/\Delta r_{CC}$ were calculated for those bonds for which structural and geometric independence could be maintained. In general, the derivatives for CC bonds in the rings were smaller than for any of the open

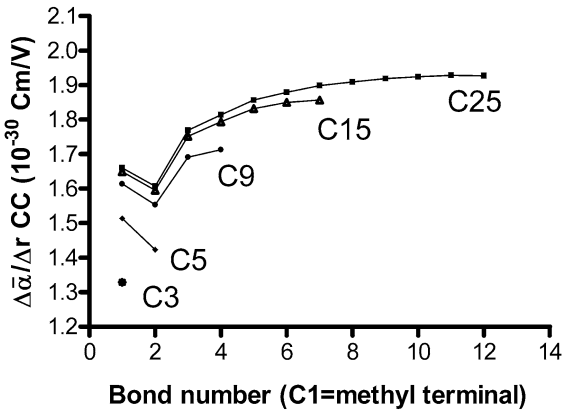


Fig. 4.8 Patterns in $\Delta\bar{\alpha}/\Delta r_{CC}$ in the straight-chain alkanes. Derivatives for each CC bond, numbered from the beginning of a chain (C1–C2 = bond 1). Data were calculated at the HF/D95(d,p) level for propane (C3), pentane (C5), nonane (C9), pentadecane (C15), and pentacosane (C25). Data taken from Ref. [21].

chains, from a low of 0.904 for the vertical bonds in hexaprismane to a high of 1.293 for the ring CC bonds in 1,4-dimethylcyclobutane. An increase in the calculated strain energies correlated well ($r^2 = 0.58$, $P < 0.0001$ for $n = 21$) with the decreasing $\Delta\bar{\alpha}/\Delta r_{CC}$ [21], but this is clearly only part of the story for these complicated structures.

4.7.4
Unsaturated Hydrocarbons and the Silanes: C–H, C=C, and Si–Si Derivatives

In keeping with the electrically conductive nature of polyenes and silanes, we observe much larger $\Delta\bar{\alpha}/\Delta r$ for these molecules [22]. The $\Delta\bar{\alpha}/\Delta r_{CH}$ for the alkenes follow the same overall patterns as for their saturated counterparts, in that the stretch of the terminal CH_{ip} bond produces the largest change in α, but the magnitude falls off much more rapidly for the CH bonds along the chain. The $\Delta\bar{\alpha}/\Delta r_{C=C}$ values are very large and increase rapidly with chain length (Fig. 4.9).

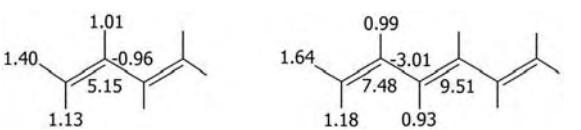


Fig. 4.9 Patterns in $\Delta\bar{\alpha}/\Delta r_{CH}$ and $\Delta\bar{\alpha}/\Delta r_{CC}$ in *trans*-1,3-butadiene and *trans*,*trans*-1,3,5-hexatriene. The end-of-chain effect is magnified by the polarizability of the π orbitals down the length of the chain, and reduced perpendicular to the chain, compared with similar values for the straight-chain alkanes. Data taken from Ref. [22].

Our silane calculations include only up to Si_5H_{12} (data not shown [22]). For $\Delta\bar{\alpha}/\Delta r_{SiH}$ and $\Delta\bar{\alpha}/\Delta r_{Si-Si}$ the pattern is the same as for the alkanes – the derivative is largest for SiH_{ip} and increases with chain length. The derivatives for the SiH_{op} bonds are similar to those along the chain and decrease with chain length.

4.8 Conclusion

Results from QTAIM analysis for the molecular polarizability derivative performed on a catalog of molecules can be summarized in two principal conclusions:

1. The elements of structure, viz. alignment (cis/trans, eq/ax, in-plane/out-of-plane), steric hindrance, and strain, are local in nature. For rings, strain applies to every atom of that ring.
2. The effects of these local features are distributed across the molecule.

In a way, QTAIM analysis of $\partial\bar{\alpha}/\partial r$ only serves to emphasize that the molecular wavefunction is a function of all nuclear and electronic degrees of freedom simultaneously.

The QTAIM atoms are as physical as the molecules from which they are recovered, arising naturally from the principles of quantum mechanics. Chemical intuition, based on experiment and theory, relies on identification of the atomic constituents of molecules as a first step, and then on application of established patterns of chemical connectivity, or bonding, to obtain principles of molecular function from structure. We, as chemists, do not work in Hilbert space, but in real space. Our building blocks are not state vectors but concrete, physical atoms. QTAIM coupled with quantum chemical calculations serves as a test bench to explore the intricacies of structure and the interplay between that structure and the resulting molecular properties. When interesting properties appear, or fail to appear, from novel structures, experiment can follow and be supported by a sound theoretical foundation. When interesting properties appear in seemingly standard structures devoid of exotic structural motifs, the need for theoretical analysis is equally clear.

Measurement of absolute Raman scattering intensities and their accurate reproduction through QTAIM are both challenging tasks. QTAIM informs our existing notions of conformation with an understanding based on the atom-by-atom contributions to the changes in molecular properties that arise from local environmental effects. It turns out for $\partial\bar{\alpha}/\partial r$ that the molecular framework itself is imprinted on the molecular response. Knowledge of the underlying atomic-level topological features of the electron density gives insight into how to design desired function into molecular systems. We are free to modify the local environments of functional groups in any way we choose, including the doping of crys-

tals. The operating principle is that a local environmental perturbation has a global effect on a molecular property.

We began this chapter with an investigation of anomalous Raman scattering intensities from CH functional groups. The first impression might have been that this was a rather esoteric probe of a mundane group of atoms. Through our QTAIM analysis we have seen that the Raman intensity is a sensitive probe of structure-imprinted, structure-mediated charge flow and rearrangement within the constraints imposed on the molecular framework by the topology of the electron density.

Authors' Note

It has become common in the literature to see the QTAIM atoms referred to as ''Bader's atoms''. It is a testament to his belief in the supremacy of experimental observation and the rigors of quantum mechanics that Bader himself decries this tendency. They are not some construct arising out of an artificial and arbitrary partitioning of space; they are *the* quantum mechanical atoms.

References

1 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, New York, **1990**.
2 T. S. Koritsanszky, P. Coppens, *Chem. Rev.* **2001**, *101*, 1583–1627.
3 G. Herzberg, *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules*, Van Nostrand Reinhold Co. Inc., New York, **1945**.
4 E. B. Wilson, J. C. Decius, P. C. Cross, *Molecular Vibrations*, McGraw–Hill, New York, **1955**.
5 G. Fogarasi and P. Pulay (Ed.: J. R. Durig), *Vibrational Spectra and Structure*, Elsevier, **1985**, 14, Chapter 3.
6 K. M. Gough, W. F. Murphy, *J. Chem. Phys.* **1987**, *87*, 1509–1519.
7 D. A. Long, *Proc. Roy. Soc. London, Ser. A.* **1952**, *217*, 203; M. Gussoni (Ed.: R. J. H. Clark, R. E. Hester), *Advances in Infrared and Raman Spectroscopy*, Heyden & Son, London, **1980**, Vol. 6.
8 K. M. Gough, W. F. Murphy, *J. Chem. Phys.* **1986**, *85*, 4290–4296.
9 K. M. Gough, W. F. Murphy, T. Stroyer-Hansen, E. Norby-Svendsen, *J. Chem. Phys.* **1987**, *87*, 3341–3345.
10 W. F. Murphy, J. M. Fernandez-Sanchez, K. Raghavachari, *J. Phys. Chem.* **1991**, *95*, 1124–1139.
11 K. M. Gough, H. K. Srivastava, *J. Phys. Chem.* **1996**, *100*, 5210–5216.
12 K. M. Gough, W. F. Murphy, *J. Mol. Struct.* **1990**, *224*, 73–88.
13 R. Dawes, K. M. Gough, *J. Chem. Phys.* **2004**, *121*, 1278–1284.
14 R. L. McCreery (Ed.: J. D. Winefordner), *Raman Spectroscopy for Chemical Analysis*, Vol. 157 in Chemical Analysis, Wiley Interscience, **2000**.
15 G. Placzek; U. S. Atomic Energy Commission, UCRL-Trans-524(L); **1962**. Translated from *Handbuch der Radiologie*, 2nd edn, Marx, E. Ed.; Akademisch: Leipzig, 1934, Vol. 6, Part II, 205–374.
16 R. G. Snyder, A. J. Aljibury, H. L. Strauss, H. L. Casal, K. M. Gough, W. F. Murphy, *J. Chem. Phys.* **1984**, *81*, 5352–5361.

17 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Chem. Phys.* **1993**, *98*, 9669–9677.

18 K. M. Gough, H. K. Srivastava, K. Belohorcová, *J. Phys. Chem.* **1994**, *94*, 771–776.

19 K. M. Gough, H. K. Srivastava, *J. Phys. Chem.* **1996**, *100*, 5210–5216.

20 K. M. Gough, J. R. Dwyer, *J. Phys. Chem. A.* **1998**, *102*, 2723–2731.

21 K. M. Gough, J. R. Dwyer, R. Dawes, *Can. J. Chem.* **2000**, *78*, 1035–1043.

22 C. Lupinetti, K. M. Gough, *J. Raman Spectroscopy.* **2002**, *33*, 147–154.

23 K. M. Gough, C. Lupinetti, R. Dawes, *J. Comput. Meth. Sci. Eng.* **2004**, *4*, 597–609.

24 R. Dawes, K. M. Gough, *J. Chem. Phys.* **2004**, *121*, 1278–1284.

25 M. J. Frisch et al., Gaussian 03, Revision A.1, Gaussian, Inc., Pittsburgh PA, 2003.

26 F. Biegler-König, AIM2000, University of Applied Sciences, Bielefeld, Germany.

27 K. M. Gough, *J. Chem. Phys.* **1989**, *91*, 2424–2432.

28 R. Dawes, Ph.D. Thesis, University of Manitoba, **2004**.

29 M. G. Papadopoulos, J. Waite, A. D. Buckingham, *J. Chem. Phys.* **1995**, *102*, 371.

30 Y. Le Duff, W. Holzer, *J. Chem. Phys.* **1974**, *60*, 2175–2178.

31 S. Montero, D. Bermejo, *Mol. Phys.* **1976**, *32*, 1229.

32 K. Kerl, H. Hausler, *Ber. Bunsenges. Phys. Chem.* **1984**, *88*, 992.

5
Topological Atom–Atom Partitioning of Molecular Exchange Energy and its Multipolar Convergence

Michel Rafat and Paul L. A. Popelier

5.1
Introduction

Without force fields the structure and dynamics of sizeable systems (biopolymers or condensed matter) would be beyond the predictive power of present day computers. The development of force fields has a long and successful history, culminating into products such as AMBER [1], CHARMm [2] or GROMOS [3], to name just a few. Despite their ubiquitous use and success they define a paradigm of accuracy, largely accepted by a community of users, despite known inherent limitations. It is very difficult to alter this paradigm and achieve a new level of accuracy unless the underlying principles of force field design are also altered, possibly drastically. That there is an increasing urgent need for sustained research effort in this direction is backed up by a recent review [4] of Ponder and Case on force fields for protein simulations. They concluded that *"An increase in computer power of at least two orders of magnitude should occur over the next decade. Without further research into the accuracy of force-field potentials, future macromolecular modeling may well be limited more by validity of the energy functions, particularly electrostatic terms, than by technical ability to perform the computations."*

There are several keys to a much needed alternative approach, all based on sound principles. The first key is replacing the intrinsically limited point charges by multipole moments. The second key is to sample as much information as possible from reduced density matrices (both 1st and 2nd order, starting with the electron density). This means that molecules and molecular complexes provide the information with which the force field is endowed, fitted, or trained. This line of attack come under the supermolecular rather than perturbation approach. The third key is a robust partitioning method to generate atomic information. Quantum chemical topology (QCT) [5, 6] is chosen, given its deep roots in quantum mechanics, its widespread use, and because it generates finite non-overlapping atoms. A fourth key, which we mention in passing, is the use of neural networks to capture the fluctuation of multipole moments in response to

a change in nuclear positions. This method deals with polarization and is currently under investigation in our laboratory.

Although the literature on the application of point-charge potentials is huge, it is recognized they have inherent limitations, for example the modeling of lone pairs and aromatic rings. Several groups (including one in Accelrys) have noticed the limitations of point charges in their work and make explicit statements [7–18] in support of multipoles. For example, in their work on the prediction of (crystal) polymorphs Day et al. [11] went as far as to assert, that "*the atomic charge calculations might have motivated an exploration of kinetic reasons why hydrogen bonded dimers are found in the crystal structure of oxazolidine-2,5-dione in preference to chains; energy calculations with the atomic multipoles model suggest that the preference is simply energetic and any exploration of other determining factors would have been ill-founded.*" Another important example is that of Batista et al. [19], who stated that "*point-charge models are typically tailored to be consistent with various properties of liquid water, but may not reproduce accurately the electric fields in other environments, such as water clusters, ice, surfaces, and interfaces.*" This insight is important for peptide solvation and docking, where small interstitial water clusters can appear.

Over the last few years we have systematically explored the behavior of multipole moments of topological atoms. In a study on the atomic partitioning of the molecular potential [20] we demonstrated the favorable convergence of the topological multipole expansion and the reason why [21]. Via a continuous multipole method involving Bessel functions it was subsequently proven [22] that the region of convergence of the potential could be enlarged. The introduction of inverse moments [23] enabled the potential to converge everywhere. The electrostatic interaction between topological atoms could also be successfully expressed as a convergent multipole expansion and could therefore be used for prediction [24] of van der Waals complexes and hydrogen-bonded DNA base pairs [25]. The convergence of the multipole expansion received attention [26] in atom–atom partitioning of intramolecular and intermolecular Coulomb energy. In that work seven systems were analyzed – an ethyne dimer, a hydrogen fluoride dimer, a water dimer, butane, 1,3,5-hexatriene, acrolein, and *cis*-urocanic acid. The current work can be seen as a counterpart of that study but for the exchange interaction rather than the Coulomb interaction. Here we focus on the same systems except the last. Here, however, the convergence analysis is extended to higher rank and includes forces rather than just energies.

As far as we know there is no force field that incorporates exchange energy, but it is possible that it will feature in a future force field. Similarly, the kinetic energy, another term in the energy partitioning of any quantum system [27], does not feature in force fields. Yet, these terms greatly affect the total energy surface of molecules and molecular assemblies and hence their structure and dynamics. Their effect is typically absorbed in the fitted bonded and non-bonded terms. In this contribution we only touch upon the question of how exchange energy may be adopted by a force field. As before, aware of the complexity of this question, we judge it to be important to first systematically analyze the convergence behavior of a high-rank multipole expansion of the exchange energy. That such an analysis

can lead to unexpected results was proven in a recent paper [28], which showed that both 1,3 and 1,4 interactions can be described on the same footing as 1,n (n > 4) interactions by a convergent multipole expansion of the Coulomb energy of the participating atom pairs.

5.2 Theoretical Background

Within the Born–Oppenheimer approximation the total closed-shell Hartree–Fock energy is given by:

$$E = -\frac{1}{2}\int d\mathbf{r}_1 \nabla_1^2 \rho(\mathbf{r}_1,\mathbf{r}_2)|_{\mathbf{r}_1=\mathbf{r}_2} + \frac{1}{2}\iint d\mathbf{r}_1\, d\mathbf{r}_2 \frac{\rho_{tot}(\mathbf{r}_1)\rho_{tot}(\mathbf{r}_2)}{r_{12}} - \frac{1}{4}\iint d\mathbf{r}_1\, d\mathbf{r}_2 \frac{\rho_1(\mathbf{r}_1,\mathbf{r}_2)\rho_1(\mathbf{r}_2,\mathbf{r}_1)}{r_{12}} \quad (1)$$

where ρ_{tot} is the total charge density, r_{12} is the inter-electron distance, and $\rho_1(\mathbf{r}_1,\mathbf{r}_2)$ the first-order reduced density matrix or one-matrix. The latter is given by:

$$\rho_1(\mathbf{r}_1,\mathbf{r}_2) = 2\sum_i \psi_i^*(\mathbf{r}_1)\psi_i(\mathbf{r}_2) \quad (2)$$

where the sum runs over molecular orbitals ψ_i. The diagonal of this density matrix is the electron density, $\rho(\mathbf{r})$, which combined with the (total) nuclear charge density yields the total charge density:

$$\rho_{tot}(\mathbf{r}) = \sum_{\mathbf{A}} \mathbf{Z_A}\delta(\mathbf{r}-\mathbf{R_A}) - \rho(\mathbf{r}) \quad (3)$$

The first term in Eq. (1) defines the kinetic energy and $|_{\mathbf{r}_1=\mathbf{r}_2}$ means the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ are kept separated until the Laplacian has operated on only one ψ, whereupon the coordinates are set equal to each other.

The total energy can be partitioned in terms of topological atoms, denoted Ω, via:

$$E = -\frac{1}{2}\sum_A \int_{\Omega_A} d\mathbf{r}_1 \nabla^2{}_1 \rho(\mathbf{r}_1,\mathbf{r}_2)|_{\mathbf{r}_1=\mathbf{r}_2} + \frac{1}{2}\sum_A\sum_B \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_{tot}(\mathbf{r}_1)\rho_{tot}(\mathbf{r}_2)}{r_{12}} - \frac{1}{4}\sum_A\sum_B \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_1(\mathbf{r}_1,\mathbf{r}_2)\rho_1(\mathbf{r}_2,\mathbf{r}_1)}{r_{12}} \quad (4)$$

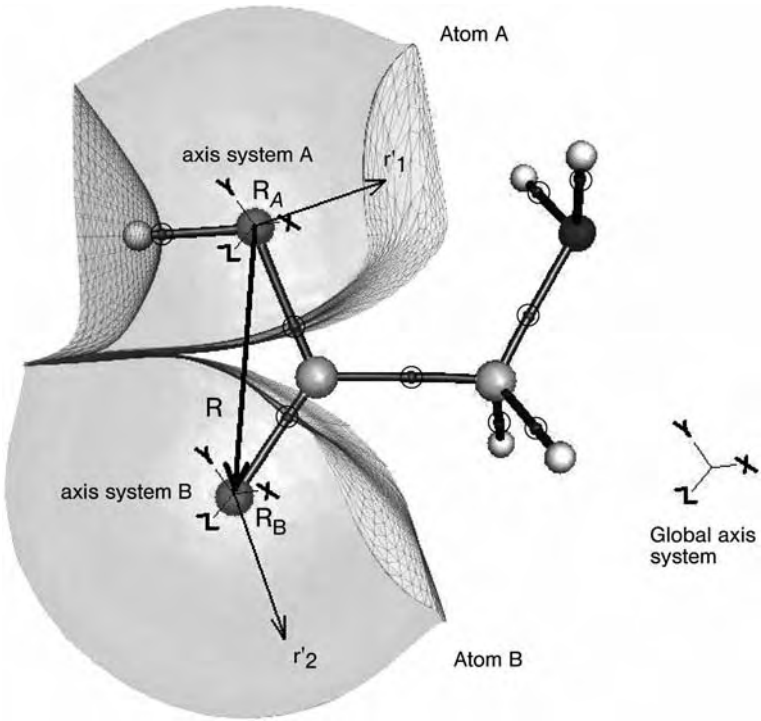


Fig. 5.1 Atomic basins of the oxygen atoms in glycine. The nuclei are shown as spheres. The bond critical points are shown by encircled dots. The interatomic surfaces are represented by small triangles. The basins are capped by an isodensity envelope at 10^{-3} au. R represents the internuclear vector, R_A and R_B the positions of the nuclei *A* and *B* in the global axis system, and r_1 and r_2 describe the electron density in the basins *A* and *B*, respectively. The global axis system is shown outside the molecule and centered on the global origin.

where the double sum $\sum_A \sum_B$ is not restricted in any way, i.e. the interaction between two different atoms (A ≠ B) is counted twice and the self-interaction of each atom (A = B) is included. The second group of terms expresses the Coulomb interaction and featured in earlier work [26]. This work focuses on the third group of terms. Figure 5.1 clarifies how the inter-electron distance r_{12} is related to the internuclear distance and the nuclear positions.

Because the molecular orbitals are real it follows that $\rho_1(\mathbf{r}_1, \mathbf{r}_2) = \rho_1(\mathbf{r}_2, \mathbf{r}_1)$ and, because $r_{12} = |\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|$, the exchange energy between atoms *A* and *B* (appearing in Eq. 4) has the form:

$$E_X^{AB} = -\int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{|\rho_1(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|} \tag{5}$$

Computation of the exchange energy in Eq. (5) involves a six-dimensional (6D) integral. The same type of integral has to be evaluated in the computation of the Coulomb interaction but there the numerator, $\rho_{\text{tot}}(\mathbf{r}_1)\rho_{\text{tot}}(\mathbf{r}_2)$, is already separated (or factorized) in terms of the variables $\mathbf{r}_1$ and $\mathbf{r}_2$. Because of the entanglement between variables $\mathbf{r}_1$ and $\mathbf{r}_2$ in the one-density matrix, the 6D integration for the exchange energy is much slower than its Coulomb counterpart. Indeed, if there are n quadrature points in each atomic grid, the numerator (i.e. the one-electron density matrix) must be evaluated for each pair of grid points, or n^2 times. In the Coulomb 6D integral, on the other hand, the charge density of each basin must be evaluated in n quadrature points, corresponding to only $2n$ evaluations of the electron density in total. In practice, the most straightforward way of separating the variables $\mathbf{r}_1$ and $\mathbf{r}_2$ is to express $\rho_1(\mathbf{r}_1, \mathbf{r}_2)$ by means of molecular orbitals:

$$E_X^{AB} = -4 \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \sum_i \sum_j \frac{\psi_i(\mathbf{r}_1)\psi_i(\mathbf{r}_2)\psi_j(\mathbf{r}_1)\psi_j(\mathbf{r}_2)}{|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|} \tag{6}$$

where i and j are summed throughout the molecular orbitals. At this stage it is convenient to define the overlap function at a point r:

$$S_{ij}(\mathbf{r}) = 2\psi_i(\mathbf{r})\psi_j(\mathbf{r}) \tag{7}$$

where we point out that the occupation number of two is absorbed in S_{ij}. Substituting Eq. (7) into Eq. (6) leads to:

$$E_X^{AB} = -\int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \sum_i \sum_j \frac{S_{ij}(\mathbf{r}_1)S_{ij}(\mathbf{r}_2)}{|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|} \tag{8}$$

This 6D integral is computationally less expensive to evaluate than that in Eq. (5), because the overlap function is computed in the two atomic grids separately, which is reminiscent of the computation of $\rho_{\text{tot}}(\mathbf{r}_1)\rho_{\text{tot}}(\mathbf{r}_2)$. The cost of Eq. (8) is still larger than the 6D integration of the Coulomb energy, however.

Both the exchange and Coulomb energy can be expressed via a multipole expansion, which has been described in detail before [26]. At the cost of possible lack of convergence this expansion separates the variables $\mathbf{R}$, $\mathbf{r}_1$, and $\mathbf{r}_2$, which are entangled in the expression $|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|^{-1}$.

A binomial Taylor expansion of $|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|^{-1}$ and subsequent application of an addition theorem for regular spherical harmonics [29, 30] factorizes the electronic $(\mathbf{r}_1, \mathbf{r}_2)$ and geometric $(\mathbf{R})$ coordinates as follows:

$$\frac{1}{|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|} = \sum_{l_1=0}^{\infty} \sum_{l_2=0}^{\infty} \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} T_{l_1 l_2 m_1 m_2}(\mathbf{R}) R_{l_1 m_1}(\mathbf{r}_1) R_{l_2 m_2}(\mathbf{r}_2) \tag{9}$$

and

$$T_{l_1 l_2 m_1 m_2}(\mathbf{R}) = (-1)^{l_1}\sqrt{\frac{(2l_1+2l_2+1)!}{(2l_1)!(2l_2)!}} \times \begin{pmatrix} l_1 & l_2 & l_1+l_2 \\ m_1 & m_2 & -(m_1+m_2) \end{pmatrix} I_{l_1+l_2,\,-(m_1+m_2)}(\mathbf{R}) \tag{10}$$

where the expression in large brackets is a Wigner 3*j* symbol and $R_{lm}(\mathbf{r})$ and $I_{lm}(\mathbf{r})$ are the regular and irregular normalized spherical harmonics $Y_{lm}(\theta,\varphi)$, respectively:

$$R_{lm}(\mathbf{r}) = \sqrt{\frac{4\pi}{2l+1}}\, r^l Y_{lm}(\theta,\varphi) \tag{11a}$$

$$I_{lm}(\mathbf{r}) = \sqrt{\frac{4\pi}{2l+1}}\, r^{-l-1} Y_{lm}(\theta,\varphi) \tag{11b}$$

The terms in Eq. (9) can be conveniently grouped according to the power of the interaction distance $R_{\mathrm{AB}}{}^{-(l_1+l_2+1)} = R_{\mathrm{AB}}{}^{-L}$ by use of Eq. (11b), where L monitors the rank of the expansion.

Substituting Eq. (9) into Eq. (8) leads to:

$$E_X^{AB} = -\sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\sum_{m_1=-l_1}^{l_1}\sum_{m_2=-l_2}^{l_2} T_{l_1 m_1 l_2 m_2}(\mathbf{R}) \sum_{ij} Q^{ij}_{l_1 m_1}(\Omega_A) Q^{ij}_{l_2 m_2}(\Omega_B) \tag{12}$$

where the *exchange moments*, $Q^{ij}_{lm}(\Omega)$, unlike their Coulomb counterpart (the familiar atomic multipole moments), explicitly depend on the molecular orbitals:

$$Q^{ij}_{lm}(\Omega) = \int_{\Omega} d\mathbf{r}\, S_{ij}(\mathbf{r}) R_{lm}(\mathbf{r}) \tag{13}$$

Note that both the overlap function S_{ij} (Eq. 7) and the electron density (deduce from Eq. 2) incorporate the occupation number. Because they share this feature with (atomic) electrostatic multipole moments, defined as $Q_{lm}(\Omega) = \int_{\Omega} d\mathbf{r}\rho(\mathbf{r}) R_{lm}(\mathbf{r})$, they have a very close and simple relationship with exchange moments, i.e. $Q_{lm}(\Omega) = \sum_i Q^{ii}_{lm}(\Omega)$.

For completeness we point out that Eq. (10) is only valid when all multipole moments are expressed relative to the global axis system. It is possible to work with moments that are expressed relative to their own local atomic axis system. In this more general approach each atomic frame can take on an arbitrary orientation. For mathematical details we refer to Section 3.3 in Stone's book [30]. In our current analysis, however, there is no need for this generalization, because we analyze single molecules and supermolecules, each expressed relative to a sin-

gle global axis system. Each atomic local frame has the same orientation as the global frame, and this is just a special case of the more general approach described above. We mention the general approach, however, because it is in this context that Hättig [31] developed his recurrence formula for computation of the interaction tensor T(**R**). This formula enables evaluation of this tensor to an arbitrarily high rank *L*. Hence we can now go beyond the limitation of L = 5 of previous work, where explicit (i.e. pre-calculated) formulae for T(**R**) were used. The evaluation of explicit formulae, as listed up to rank L = 6 [30, 32], is more rapid, however.

The exact exchange force is defined by Eq. (14).

$$\mathbf{F}_X^{AB} = -4\int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \sum_i \sum_j \frac{\psi_i(\mathbf{r}_1)\psi_i(\mathbf{r}_2)\psi_j(\mathbf{r}_1)\psi_j(\mathbf{r}_2)}{r_{12}^3}\mathbf{r}_{12} \tag{14}$$

It is straightforward to differentiate Eq. (12) to obtain the *k*th force component from the exchange energy:

$$F_{X,k}^{AB} = \sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\sum_{m_1=-l_1}^{l_1}\sum_{m_2=-l_2}^{l_2} \frac{\partial T_{l_1 m_1 l_2 m_2}(\mathbf{R})}{\partial R_{A,k}} \sum_{ij} Q_{l_1 m_1}^{ij}(\Omega_A) Q_{l_2 m_2}^{ij}(\Omega_B) \tag{15}$$

where $R_{A,k}$ is the *k*th component of the position vector of the nucleus *A* and where we have assumed that the exchange moments do not vary with a change in nuclear positions. Note that Hättig's recursive formulae are easy to differentiate with respect to nuclear components.

The development so far has focused on (closed-shell) Hartree–Fock wave functions, but it is possible to extend the multipole formalism to post Hartree–Fock wave functions. The total (electronic) interaction energy between two topological atoms is given by Eq. (16):

$$E_{ee}^{AB} = \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{R} + \mathbf{r}_2 - \mathbf{r}_1|} \tag{16}$$

where the second-order reduced density matrix $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ can be written in terms of molecular orbitals as follows:

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \sum_{ijkl} C_{ijkl}\psi_i(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\psi_k(\mathbf{r}_2)\psi_l(\mathbf{r}_2) \tag{17}$$

The only nonvanishing *C* coefficients at the Hartree–Fock level are $C_{iijj} = 2$ and $C_{ijji} = -1$. Equation (6) is reproduced by inserting Eq. (17) into Eq. (16) and collecting the terms in the quadruple sum of Eq. (17) corresponding to $C_{ijji} = -1$. The multipole expansion for post-Hartree–Fock wave functions is generalized in a straightforward manner in Eq. (18).

$$E_{ee}^{AB} = \sum_{l_1=0}^{\infty}\sum_{l_2=0}^{\infty}\sum_{m_1=-l_1}^{l_1}\sum_{m_2=-l_2}^{l_2} T_{l_1 m_1 l_2 m_2}(\mathbf{R}) \sum_{ijkl} C_{ijkl} Q_{l_1 m_1}^{ij}(\Omega_A) Q_{l_2 m_2}^{kl}(\Omega_B) \qquad (18)$$

We note that Pendas et al. [33] developed a monadic diagonalization of $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$ to avoid the computational overhead because of the extra four summations occurring in Eqs (17) and (18). Using CI wave functions Fradera et al. [34] and Poater et al. [35] calculated electron localization and delocalization indices for topological atoms. Their indices are functions of $S_{ij}(\Omega)$, which can be identified with a ''monopolar exchange moment'' ($l = m = 0$), because from Eq. (13) we learn that $Q_{oo}^{ij}(\Omega) = \int_\Omega d\mathbf{r} S_{ij}(\mathbf{r})$, because $R_{00}(r) = 1$. In their work the C_{ijkl} coefficients were extracted from the ab initio code GAMESS, an involved procedure that was avoided by Wang and Werstiuk [36]. Taking advantage of the so called Z-vector method [37] they developed a simpler way to include the Coulomb correlation effects. By introducing natural molecular orbitals and non-integer occupation numbers the equations developed at the Hartree–Fock level can be retained for post-Hartree–Fock wave functions.

Because, in this study, we used B3LYP wave functions we need to comment on the validity and re-interpretation of the multipole formalism, especially Eq. (12) and some equations leading up to it. There are inherent differences between the exchange-correlation of the Kohn–Sham formalism and Hartree–Fock formalism [38, 39]. With Kohn–Sham density-functional theory the total electronic energy of the real, fully-interacting system is expressed as:

$$E_{tot} = T_0 + \int d\mathbf{r} \rho V_{nuc} + \frac{1}{2}\iint d\mathbf{r}_1\, d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} + E_{XC} \qquad (19)$$

where T_0 is the kinetic energy of the non-interacting reference system, and the second and third terms are the nuclear interaction energy and the Coulomb energy, respectively. The last term, E_{XC}, actually defines the Kohn–Sham exchange-correlation energy. It contains, buried within it, all the details of two-body exchange and a kinetic-energy component. Inserting Kohn–Sham orbitals in Eq. (8) to compute and analyze what looks like an ''exchange'' energy can be justified on the grounds of the generality of $\rho_1(\mathbf{r}_1, \mathbf{r}_2)$ in Eq. (5) and the similarity between the Kohn–Sham scheme and the Hartree–Fock scheme. The essence of this work is to investigate the convergence of multipole expansions of a substantial contribution to the total energy, which is not the Coulomb energy or the exact kinetic energy. For easy reference we will refer to this energy contribution simply as the ''exchange energy''.

5.3 Details of Calculations

We have modified the program MORPHY01 to integrate the exchange moments Q_{lm}^{ij} for a pair of molecular orbitals (i, j). The moments are written in an output

file. The CPU time required for the integration limited the maximum rank of the moments computed to L = 10. The atomic basins are capped at the $\rho = 10^{-7}$ a.u. isodensity envelope. Because these calculations are computationally demanding we used the default grid of MORPHY01 for the 3D integration of the moments, the number of radial and angular quadrature points being set at $(n_r, n_\theta, n_\varphi) = (80, 30, 50)$ for the β sphere and $(70, 50, 80)$ for the rest of the atomic basin. The interaction tensor T(**R**), and its derivatives, were calculated up to an arbitrary rank L by use of an in-house program containing Hättig's recurrence formulae. MORPHY01 was previously used [26] to obtain the (''exact'', i.e. non-expanded) Coulomb energy between two topological atoms. This program was modified to compute the exact exchange energy and force for two topological atoms by 6D numerical integration.

Figure 5.2 shows the geometry and numbering scheme of the six systems studied. All molecules were optimized, without imposed symmetry, by the program GAUSSIAN03 [40] at the same of level of theory, B3LYP/6-311 + G(2d,p), as in the earlier counterpart study [26] on the Coulomb interaction.

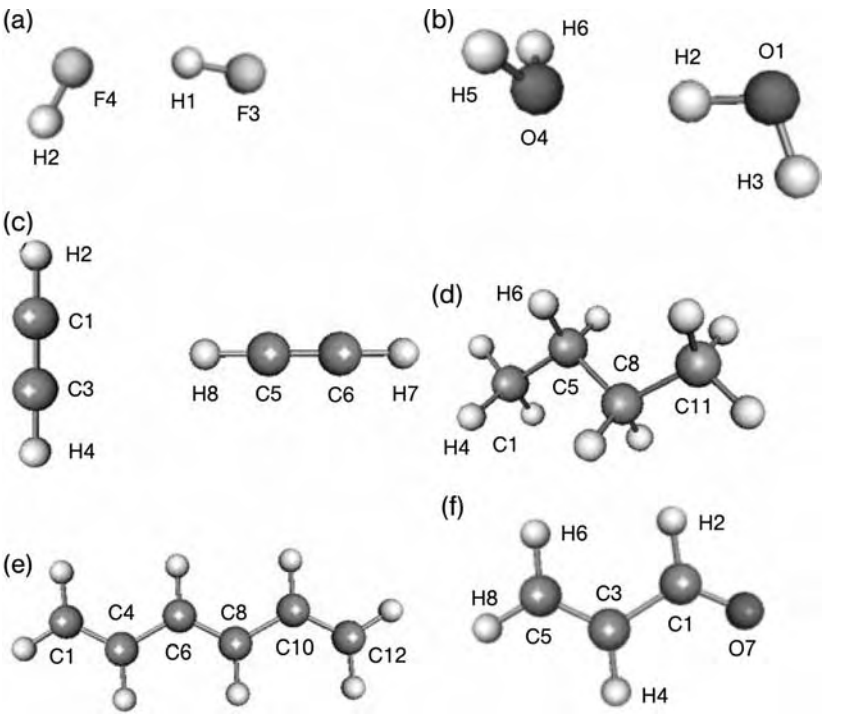


Fig. 5.2 Geometry and numbering scheme of the six systems studied: (a) HF dimer, (b) H_2O dimer, (c) C_2H_2 dimer, (d) butane, (e) 1,3,5-hexatriene, (f) acrolein.

Table 5.1 Exchange energy (kJ mol^{-1}) between atomic basins with increasing nuclear separation R in $(HF)_2$ from multipole expansion (increasing rank L) and 6D numerical integration (exact).

R (a.u.)	H1–F3[a] 1.76	H2–F4[a] 1.75	H1–F4 3.44	H1–H2 4.49	F3–F4 5.17	H2–F3 6.02
L = 1	−660.7	−733.0	−44.0	−0.4	−35.6	−0.4
L = 2	−723.6	−747.9	−61.3	−0.4	−46.9	−0.3
L = 3	−791.1	−858.6	−68.0	−0.4	−51.1	−0.3
L = 4	−784.2	−822.1	−70.1	−0.4	−52.1	−0.3
L = 5	−702.0	−725.7	−70.2	−0.4	−52.5	−0.3
L = 6	−694.0	−773.7	−69.5	−0.4	−52.7	−0.3
L = 7	−770.3	−742.1	−68.9	−0.4	−52.7	−0.3
L = 8	−759.4	−699.4	−68.8	−0.4	−52.5	−0.3
L = 9	−672.4	−1272.6	−69.1	−0.4	−52.2	−0.3
L = 10	−823.4	−791.2	−69.6	−0.4	−52.2	−0.3
Exact	−740.4	−783.6	−69.8	−0.4	−52.4	−0.3

[a] Note that in this and subsequent tables the bonded interactions (i.e. between two atoms separated by one or two bonds, or 1,2 and 1,3 interactions) are colored in light grey.

5.4 Results and Discussion

5.4.1 Convergence of the Exchange Energy

The HF dimer often appears as a simple prototype system featuring medium to strong hydrogen bonding between polar molecules. Table 5.1 shows all possible atom–atom interactions ($\binom{4}{2} = 6$). It is clear that the multipole expansion does not converge for the nearly equivalent covalently bonded atoms (H1–F3 and H2–F4), as is observed for the corresponding Coulomb interaction. It is pleasing to learn that the hydrogen bonded interaction (H1–F4) stabilizes rapidly toward the exact value of −69.8 kJ mol^{-1}, already at L = 4. This was also observed for the hydrogen bond Coulomb interaction. Perhaps surprising is the rather sizeable (but properly converged) exchange energy between the fluorine atoms.

The exact and multipole expanded exchange energies of eleven relevant interactions in the water dimer are presented in Table 5.2. There is a total of $\binom{6}{2} = 15$ interactions but four can be left out because of (near) planar symmetry. The pattern observed here is very similar to that of the HF dimer. As expected, the multipole expansion is unable to converge for bonded interactions. For the non-bonded

Table 5.2 Exchange energy (kJ mol^{-1}) between atomic basins with increasing nuclear separation R in $(H_2O)_2$ from multipole expansion (increasing rank L) and 6D numerical integration (exact). The labels correspond to those in Fig. 5.2.

R (a.u.)	O4–H5 1.82	O1–H3 1.82	O1–H2 1.84	H2–H3 2.90	H5–H6 2.90	H2–O4 3.66	H2–H5 4.62
L = 1	−1009.0	−1064.8	−870.8	−7.5	−8.4	−55.3	−0.7
L = 2	−1087.6	−1135.9	−1002.5	−4.9	−5.1	−78.6	−0.7
L = 3	−1214.7	−1269.6	−1080.4	−5.8	−6.5	−87.6	−0.6
L = 4	−1170.3	−1232.6	−1062.6	−5.6	−6.0	−90.1	−0.6
L = 5	−994.6	−1021.9	−940.2	−5.5	−6.1	−89.6	−0.6
L = 6	−1060.3	−1098.5	−905.2	−5.6	−6.2	−88.3	−0.6
L = 7	−1151.3	−1274.7	−1072.9	−5.6	−6.1	−87.4	−0.6
L = 8	−944.3	−917.7	−1170.6	−5.6	−6.1	−87.6	−0.6
L = 9	−2009.2	−2171.8	−763.8	−5.6	−6.1	−88.4	−0.6
L = 10	−1565.1	−2265.0	−702.7	−5.7	−6.2	−88.9	−0.6
Exact	−1109.6	−1161.1	−1003.7	−5.6	−6.1	−88.2	−0.7

R (a.u.)	O1–O4 5.49	O1–H5 6.31	H3–O4 6.34	H3–H5 7.38
L = 1	−34.6	−0.5	−0.8	0.0
L = 2	−44.8	−0.5	−0.8	0.0
L = 3	−48.4	−0.5	−0.8	0.0
L = 4	−48.8	−0.5	−0.8	0.0
L = 5	−49.1	−0.5	−0.8	0.0
L = 6	−49.3	−0.5	−0.8	0.0
L = 7	−49.3	−0.5	−0.8	0.0
L = 8	−49.0	−0.5	−0.8	0.0
L = 9	−48.8	−0.5	−0.8	0.0
L = 10	−48.8	−0.5	−0.8	0.0
Exact	−49.0	−0.5	−0.8	0.0

interactions, however, the expansions predict a value very close to that from 6D integration. Even at short range (for instance the H2–H3 and H5–H6 interactions), for which the Coulomb expansion is expected to fail, the exchange interaction has an error of less than 0.1 kJ mol^{-1}. There are only small differences (less than 1 kJ mol^{-1}) between the expansions at rank L = 6 and L = 10 for non-bonded interactions. Although the bonded O–H interaction largely dominates the other interactions, the intermolecular interactions between H2 and O4, which represents the hydrogen bond, and the oxygen–oxygen interactions are not negli-

Table 5.3 Exchange energy (kJ mol^{-1}) between selected atomic basins (AM–BN, where M < N; Fig. 5.2) in $(C_2H_2)_2$ from multipole expansion (increasing rank *L*) and 6D numerical integration (exact).

R (a.u.)	H2–C3 4.27	H2–H4 6.28	H2–C5 8.03	H2–H8 6.23	C3–H4 2.01	C3–H5 7.48	C3–C6 9.72	C3–H8 5.49
L = 1	−45.9	−1.5	−0.1	−0.3	−1252.9	−3.1	−0.4	−10.2
L = 2	−46.5	−1.3	−0.1	−0.4	−1638.9	−4.2	−0.5	−14.1
L = 3	−42.1	−1.3	−0.1	−0.4	−1600.4	−4.6	−0.5	−15.6
L = 4	−42.8	−1.3	−0.1	−0.4	−1455.8	−4.7	−0.5	−16.2
L = 5	−44.2	−1.3	−0.1	−0.4	−1366.9	−4.8	−0.5	−16.3
L = 6	−43.2	−1.3	−0.1	−0.4	−1398.6	−4.8	−0.5	−16.3
L = 7	−42.8	−1.3	−0.1	−0.4	−1665.5	−4.8	−0.5	−16.2
L = 8	−43.7	−1.3	−0.1	−0.4	−1973.7	−4.8	−0.5	−16.2
L = 9	−43.5	−1.3	−0.1	−0.4	−1323.3	−4.8	−0.5	−16.2
L = 10	−42.9	−1.3	−0.1	−0.4	117.4	−4.8	−0.5	−16.3
Exact	−43.3	−1.3	−0.1	−0.4	−1486.5	−4.8	−0.5	−16.3

R (a.u.)	H4–C5 8.05	C5–C6 2.26	C5–H7 4.27	C5–H8 2.01	C6–H7 2.01	C6–H8 4.27	H7–H8 6.28
L = 1	−0.1	−3317.1	−47.1	−1222.7	−1260.6	−44.1	−1.4
L = 2	−0.1	−4062.3	−47.5	−1621.8	−1646.8	−45.2	−1.3
L = 3	−0.1	−3591.0	−42.8	−1580.8	−1605.7	−40.8	−1.3
L = 4	−0.1	−3782.0	−43.7	−1428.4	−1461.3	−41.2	−1.3
L = 5	−0.1	−5041.4	−45.2	−1358.2	−1372.7	−42.9	−1.3
L = 6	−0.1	−4568.9	−44.1	−1398.5	−1399.6	−42.1	−1.3
L = 7	−0.1	−761.7	−43.6	−1578.5	−1689.4	−41.1	−1.3
L = 8	−0.1	−1668.4	−44.8	−1940.6	−2049.5	−42.2	−1.3
L = 9	−0.1	−18659.7	−44.5	−1568.4	−1264.1	−43.0	−1.3
L = 10	−0.1	−23057.4	−43.6	−1528.1	466.0	−41.3	−1.3
Exact	−0.1	−4003.1	−44.2	−1465.9	−1493.0	−42.0	−1.3

gible. It is, again, gratifying to note that the hydrogen-bond exchange energy converges rapidly.

The exact and multipole-expanded exchange energies of fifteen selected atom–atom interactions in the ethyne dimer are presented in Table 5.3. The covalent bonds C–H and C≡C do not converge, the exchange energy of the latter bond being roughly three times that of the former. The next largest values are approximately a factor of 30 smaller and occur between the non-bonded C and H atoms of the respective ethyne moieties. Very low rank expansions approximate the exact

Table 5.4 Exchange energy (kJ mol^{-1}) between selected atomic basins (AM–BN, where M < N; Fig. 5.2) in butane from multipole expansion (increasing rank *L*) and 6D numerical integration (exact).

R (a.u.)	C1–H4 2.07	C1–C5 2.89	C1–H6 4.08	C1–C8 4.83	C1–C11 7.42	H4–C11 8.03	C5–H6 2.07	C5–C8 2.89
L = 1	−1219.3	−906.4	−27.8	−27.0	−4.2	−0.3	−1192.5	−888.7
L = 2	−1531.9	−1332.6	−31.5	−34.9	−4.7	−0.4	−1509.1	−1312.5
L = 3	−1598.9	−1546.2	−33.2	−38.1	−4.8	−0.4	−1579.9	−1526.6
L = 4	−1478.5	−1603.0	−33.3	−39.5	−4.8	−0.4	−1476.2	−1591.1
L = 5	−1363.2	−1606.1	−33.5	−40.2	−4.8	−0.4	−1368.3	−1594.2
L = 6	−1356.8	−1578.5	−33.4	−40.4	−4.8	−0.4	−1358.1	−1568.1
L = 7	−1531.2	−1561.0	−32.9	−40.5	−4.8	−0.4	−1456.9	−1545.0
L = 8	−1731.5	−1533.3	−33.1	−40.4	−4.8	−0.4	−1588.1	−1535.7
L = 9	−1569.5	−1551.1	−33.4	−40.3	−4.8	−0.4	−1626.0	−1532.0
L = 10	−1446.1	−1579.7	−33.5	−40.2	−4.8	−0.4	−1515.3	−1577.3
Exact	−1463.5	−1568.1	−33.2	−40.4	−4.8	−0.3	−1449.6	−1556.0

exchange energies quite well. Several interactions, occurring between atoms more than 10 a.u. apart, for example H2–C6, H2–H7, C3–H7, and H4–C6 have vanishing exchange energy (i.e. <0.1 kJ mol^{-1}). The convergent behavior of the expansion at short range is excellent, because the difference of energy between ranks L = 9 and L = 10 is always less than 0.01 kJ mol^{-1} for non-bonded interactions and the maximum difference between L = 6 and L = 10 is 0.13 kJ mol^{-1}. The nonbonded interactions, for example C3–H8 and C3–C5, are not as strong as in the water dimer but are still significant in terms of "chemical accuracy" (1 kcal mol^{-1} or 4 kJ mol^{-1}).

The exact and multipole-expanded exchange energies of eight selected atom–atom interactions in butane are shown in Table 5.4. As in the previous systems the convergence is faster than for the Coulomb interactions. Remarkably, for some bonded interactions (for example C5–C8) the difference between the exact and L = 10 is in the region of 20 kJ mol^{-1}. Also, the exchange energy between C and H in covalent C–H bonds is fairly constant, whether in butane or ethyne.

The exact and multipole-expanded exchange energies of ten selected atom–atom interactions in 1,3,5-hexatriene are shown in Table 5.5. In this conjugated system the exchange energy can still be quite large over long distances. For example the exchange energy between the furthest atoms, C1 and C12, is −7.9 kJ mol^{-1}. It is not clear why this value is larger in absolute magnitude than that for C1–C10, and appears as an exception in a regularly decreasing series of absolute energy values, starting with C1–C4. The same anomaly was found in the Coulomb interaction [26]. The formally CC double bonds, C1–C4 and C6–C8 are both associated with an energy of approximately -2.5×10^3 kJ mol^{-1}. A single (sigma) CC bond, as found in butane (C1–C5, C5–C8), corresponds to

Table 5.5 Exchange energy (kJ mol^{-1}) between selected carbon basins (CM–CN, where M < N; Fig. 5.2) in 1,3,5-hexatriene from multipole expansion (increasing rank *L*) and 6D numerical integration (exact).

R (a.u.)	C1–C4 2.53	C1–C6 4.66	C1–C8 6.99	C1–C10 9.32	C1–C12 11.60
L = 1	−1822.8	−44.4	−36.2	−2.3	−8.2
L = 2	−2458.4	−54.0	−36.1	−2.3	−8.0
L = 3	−2513.6	−55.0	−34.6	−2.3	−7.9
L = 4	−2536.0	−54.8	−34.7	−2.3	−7.9
L = 5	−2798.9	−55.2	−34.9	−2.3	−7.9
L = 6	−2916.1	−56.1	−34.9	−2.3	−7.9
L = 7	−2222.2	−56.7	−34.8	−2.3	−7.9
L = 8	−1346.2	−55.7	−34.8	−2.3	−7.9
L = 9	−3869.5	−54.2	−34.9	−2.3	−7.9
L = 10	−9335.7	−55.7	−34.8	−2.3	−7.9
Exact	−2605.6	−55.5	−34.9	−2.3	−7.9

R (a.u.)	C4–C6 2.73	C4–C8 4.67	C4–C10 7.21	C4–C12 9.32	C6–C8 2.54
L = 1	−1102.9	−41.8	−4.9	−2.3	−1695.1
L = 2	−1625.7	−51.7	−5.4	−2.3	−2332.4
L = 3	−1821.3	−52.9	−5.4	−2.3	−2431.9
L = 4	−1863.8	−52.6	−5.4	−2.3	−2448.6
L = 5	−1892.1	−53.0	−5.4	−2.3	−2635.8
L = 6	−1909.8	−54.0	−5.4	−2.3	−2758.6
L = 7	−1803.0	−54.4	−5.4	−2.3	−2282.2
L = 8	−1599.5	−53.6	−5.4	−2.3	−1542.5
L = 9	−1816.2	−52.5	−5.4	−2.3	−3031.5
L = 10	−2759.3	−53.6	−5.4	−2.3	−7075.1
Exact	−1842.1	−53.9	−5.4	−2.3	−2498.6

approximately -1.5×10^3 kJ mol^{-1}. The C4–C6 (hexatriene) exchange energy of -1.8×10^3 kJ mol^{-1} could then be interpreted as, essentially, a single CC bond with some double bond character. We note that the Coulomb energy [24] of C1–C5 in butane is almost half that of C4–C6 in hexatriene. Hence it seems that the conjugation has much less effect on the exchange energy than on the Coulomb energy.

The exact and multipole-expanded exchange energies of all $\binom{8}{2} = 28$ atom–atom interactions in acrolein are shown in Table 5.6. Overall, excellent conver-

Table 5.6 Exchange energy (kJ mol^{-1}) between all basins (AM–BN, where M < N; Fig. 5.2) in acrolein from multipole expansion (increasing rank *L*) and 6D numerical integration (exact).

R (a.u.)	C1–H2 2.10	C1–C3 2.78	C1–H4 4.13	C1–C5 4.61	C1–H6 5.04	C1–O7 2.29	C1–H8 6.51	H2–C3 4.13	H2–H4 5.91	H2–C5 4.97
L = 1	−1095.3	−974.1	−22.2	−57.6	−7.3	−1692.7	−5.0	−34.6	−4.8	−7.5
L = 2	−1461.1	−1457.7	−23.0	−72.3	−8.9	−2145.5	−5.1	−35.1	−3.8	−9.8
L = 3	−1486.4	−1659.2	−23.1	−72.1	−9.2	−2152.0	−5.1	−34.7	−4.0	−10.5
L = 4	−1377.5	−1717.7	−23.2	−70.8	−9.3	−2184.3	−5.1	−34.6	−4.0	−10.8
L = 5	−1300.6	−1731.6	−23.5	−72.2	−9.4	−2380.8	−5.1	−35.2	−4.0	−10.9
L = 6	−1337.3	−1718.2	−23.4	−73.9	−9.5	−2464.4	−5.1	−35.1	−4.0	−10.9
L = 7	−1463.9	−1651.5	−23.2	−73.6	−9.5	−2132.3	−5.1	−34.7	−4.0	−10.9
L = 8	−1496.8	−1583.0	−23.3	−72.2	−9.4	−1659.0	−5.1	−34.9	−4.0	−10.9
L = 9	−1226.7	−1695.7	−23.4	−72.3	−9.5	−2419.9	−5.1	−35.3	−4.0	−11.0
L = 10	−953.3	−2001.9	−23.3	−73.9	−9.5	−4760.9	−5.1	−34.9	−4.0	−11.0
Exact	−1377.5	−1687.8	−23.1	−73.1	−9.6	−2264.9	−5.1	−37.7	−4.0	−11.0

R (a.u.)	H2–H6 4.50	H2–O7 3.81	H2–H8 7.02	C3–H4 2.05	C3–C5 2.52	C3–H6 3.98	C3–O7 4.48	C3–H8 4.00	H4–C5 4.01
L = 1	−7.1	−81.1	−0.5	−1205.7	−1867.3	−38.9	−80.1	−38.1	−36.0
L = 2	−9.9	−78.7	−0.6	−1563.5	−2494.7	−39.3	−97.6	−39.6	−37.8
L = 3	−10.8	−83.0	−0.6	−1586.7	−2542.7	−37.4	−102.9	−37.4	−36.2
L = 4	−10.9	−82.2	−0.6	−1466.4	−2578.3	−37.8	−104.3	−37.7	−36.3
L = 5	−10.9	−82.6	−0.6	−1354.3	−2848.0	−39.2	−105.0	−39.0	−37.4
L = 6	−10.9	−83.3	−0.6	−1366.2	−2926.2	−38.6	−106.1	−38.5	−36.9
L = 7	−10.9	−82.1	−0.6	−1594.6	−2255.1	−37.8	−107.0	−37.5	−36.1
L = 8	−10.9	−82.4	−0.6	−1761.7	−1521.3	−38.6	−106.1	−38.2	−36.8
L = 9	−10.9	−83.3	−0.6	−1340.7	−3844.3	−39.1	−104.6	−39.1	−37.5
L = 10	−10.9	−82.9	−0.6	−1015.4	−8410.6	−37.8	−105.4	−38.0	−36.3
Exact	−10.9	−82.6	−0.6	−1462.0	−2646.6	−42.7	−106.0	−38.4	−36.7

gence rates are observed for acrolein. For all the nonbonded interactions there is a difference between L = 9 and L = 10 of less than 2 kJ mol^{-1}. The C1–C3 bond has an exchange energy equivalent to that of a single CC bond in butane and the π bond C5–C3 is also equivalent in energy to the C1–C4 double bond in 1,3,5-hexatriene. The C=O bond (C1–O7) has a similar exchange energy.

It is possible to establish a link with the delocalization index $\delta(A, B)$, which is interpreted as a bond order. One can prove that $\delta(A, B)$ is proportional to $-E_X^{AB}R$, where R is the internuclear distance, if E_X^{AB} is restricted to L = 1 (hence $\ell_1 + \ell_2 + 1 = 1$ and therefore $\ell_1 = \ell_2 = m_1 = m_2 = 0$).

Table 5.6 *(continued)*

R (a.u.)	H4–H6 5.83	H4–O7 4.98	H4–H8 4.71	C5–H6 2.05	C5–O7 6.72	C5–H8 2.05	H6–O7 7.33	H6–H8 3.49	O7–H8 8.46
L = 1	−6.3	−8.6	−3.8	−1223.3	−36.7	−1236.8	−2.3	−25.2	−2.0
L = 2	−5.2	−11.0	−3.7	−1576.5	−36.4	−1592.6	−2.4	−19.5	−1.9
L = 3	−5.4	−11.7	−3.7	−1601.0	−35.4	−1610.8	−2.3	−20.8	−1.8
L = 4	−5.4	−11.8	−3.7	−1474.8	−35.5	−1483.9	−2.3	−21.1	−1.8
L = 5	−5.4	−11.9	−3.7	−1363.7	−35.6	−1365.3	−2.3	−20.6	−1.8
L = 6	−5.4	−11.9	−3.7	−1375.0	−35.5	−1381.4	−2.3	−20.6	−1.8
L = 7	−5.4	−11.9	−3.7	−1641.6	−35.5	−1672.5	−2.3	−20.9	−1.8
L = 8	−5.4	−11.9	−3.7	−1837.3	−35.5	−1867.2	−2.3	−20.8	−1.8
L = 9	−5.4	−11.9	−3.7	−1173.7	−35.5	−1229.5	−2.3	−20.4	−1.8
L = 10	−5.4	−11.9	−3.7	−695.0	−35.5	−807.6	−2.3	−20.9	−1.8
Exact	−5.37	−11.9	−3.7	−1475.2	−35.5	−1485.6	−2.3	−20.7	−1.8

5.4.2
Convergence of the Exchange Force

We computed the multipole expansion of the exchange force and compared it with the exact exchange force. Table 5.7 shows the results for some selected interactions of the water dimer. These two tables indicate that the convergence of the exchange force is not as good as the convergence of the exchange energy. For example, the average relative difference in the interactions of acrolein (Table 5.8) considered is 2.7% for the magnitude of the force whereas the relative difference of the energy for the same interactions is only 0.6%. In the water dimer the relative difference reaches 7.6% for the force but only 1.0% for the energy.

5.4.3
Diagonalization of a Matrix of Exchange Moments

Coulomb multipole moments depend only on the electron density, but exchange moments explicitly refer to a set of molecular orbitals. The choice of these orbitals is arbitrary, up to a unitary transformation. This is why it is interesting to investigate the intrinsic information stored in the exchange moment matrix, $\mathbf{Q^{lm}}$, which is defined by the elements $\{Q_{ij}{}^{lm}; i = 1, N; j = 1, N\}$ where N is the number of molecular orbitals. We set out to diagonalize such a matrix and inspect its eigenvalues. For a given moment the corresponding matrix $\mathbf{Q^{lm}}$ can be written:

$$Q_{ij}^{lm} = \sum_{k=1}^{N} U_{ik}^{lm} q_k^{lm} U_{jk}^{lm} \tag{20}$$

Table 5.7 Magnitude[a] of the exchange force between selected atoms (AM–BN, where M < N; Fig. 5.2) for the water dimer computed from the multipole expansion and 6D integration.

R (a.u.)	O1–O4 5.49	O1–H5 6.31	H2–H3 2.90	H2–O4 3.66	H2–H5 4.62	H3–O4 6.34	H5–H6 2.90
L = 1	11.9	0.2	4.9	28.6	0.3	0.2	5.5
L = 2	18.9	0.1	1.5	52.6	0.3	0.3	1.2
L = 3	22.6	0.2	3.2	66.6	0.2	0.3	3.8
L = 4	23.3	0.1	2.8	71.7	0.2	0.2	2.7
L = 5	23.7	0.1	2.5	70.5	0.3	0.2	2.7
L = 6	24.1	0.1	2.9	66.4	0.3	0.3	3.2
L = 7	24.2	0.1	2.8	63.3	0.2	0.3	2.7
L = 8	23.5	0.1	2.6	63.9	0.3	0.2	2.9
L = 9	22.6	0.1	2.5	67.5	0.3	0.2	2.9
L = 10	22.7	0.1	3.4	70.2	0.3	0.3	3.3
Exact	23.4	0.1	2.7	66.6	0.3	0.2	2.9

[a] The magnitude of the force between the atoms H3 and H5 (R = 7.38 au) is not shown because it is lower than 0.01 kJ mol^{-1} Å.

Table 5.8 Magnitude of the exchange force between selected atoms (AM–BN, where M < N; Fig. 5.2) for acrolein computed from the multipole expansion and 6D integration.

R (a.u.)	C1–H4 4.13	C1–C5 4.61	C1–H6 5.04	C1–H8 6.51	H2–H4 5.91	H2–H6 4.50	H2–O7 3.81	H2–H8 7.02	C3–O7 4.48	C3–H8 4.00
L = 1	10.2	23.6	2.7	1.4	1.5	3.0	40.2	0.1	33.8	18.0
L = 2	11.0	35.6	4.0	1.5	0.9	5.4	37.8	0.2	48.5	19.4
L = 3	11.1	35.4	4.3	1.5	1.0	6.4	44.2	0.2	55.2	16.3
L = 4	11.1	33.3	4.4	1.5	1.1	6.6	42.6	0.2	57.5	17.0
L = 5	11.9	36.0	4.7	1.5	1.0	6.6	43.6	0.2	59.0	19.9
L = 6	11.7	40.3	4.8	1.5	1.0	6.6	45.6	0.2	61.9	18.4
L = 7	11.1	39.5	4.7	1.5	1.0	6.7	41.5	0.2	64.5	15.2
L = 8	11.3	34.9	4.7	1.5	1.0	6.7	42.7	0.2	61.6	17.9
L = 9	11.9	35.1	4.8	1.5	1.0	6.6	46.6	0.2	55.8	21.9
L = 10	11.5	41.9	4.8	1.5	1.0	6.6	44.6	0.2	59.3	16.8
Exact	11.2	37.8	4.9	1.5	1.0	6.7	43.5	0.2	60.9	17.9

Table 5.9 Eigenvalues of the matrix of exchange moments $\mathbf{Q}^{00}$ for all atoms in the water dimer (labels as in Fig. 5.2).

n[a]	O1	H2	H3	O4	H5	H6
1	−2.0000	−0.3725	−0.4344	−2.0000	−0.4097	−0.4103
2	−1.9866	−0.0182	−0.0181	−1.9795	−0.0167	−0.0168
3	−1.9632	−0.0151	−0.0125	−1.9315	−0.0109	−0.0109
4	−1.6494	−0.0107	−0.0034	−1.6315	−0.0028	−0.0028
5	−1.5254	−0.0010	0.0000	−1.5262	−0.0001	−0.0001
6	−0.0085	−0.0005	0.0000	−0.0223	0.0000	0.0000
7	−0.0014	−0.0002	0.0000	−0.0039	0.0000	0.0000
8	−0.0005	0.0000	0.0000	−0.0020	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Population	9.1350	0.4182	0.4685	9.0970	0.4402	0.4409

[a] The number of the molecular orbital (accommodating 10 × 2 electrons in total), ranked from lowest to highest energy.

where $\{q_k{}^{lm}; k = 1, N\}$ are the eigenvalues of $\mathbf{Q}^{\mathbf{lm}}$ and the matrix $\mathbf{U}^{\mathbf{lm}}$ represents the eigenvectors. Because the trace of the matrix is invariant with regard to the unitary transformation in Eq. (20), the trace of $\mathbf{Q}^{\mathbf{lm}}$ is equal to the sum of the eigenvalues.

Table 5.9 shows an example of such "exchange eigenvalues" q_k^{lm} for each of the ten molecular orbitals and for each atom in the water dimer, when $l = m = 0$. An interesting property of this matrix is that its trace (i.e. the sum of its diagonal elements) is equal to the electrostatic multipole moments. This is illustrated for the water dimer, where the $q_k{}^{00}$ eigenvalues (of a given atom) summed over all molecular orbitals yield the population of that atom. The population of the water dimer amounts to 19.9998, illustrating the accuracy of the integration.

Using Eq. (20) we can now express the exchange energy with the diagonalized exchange moments:

$$E_X = -\sum_{l_1 m_1} \sum_{l_2 m_2} T_{l_1 m_1 l_2 m_2} \sum_{i=1}^{N} \sum_{j=1}^{N} q_i^{l_1 m_1} q_j^{l_2 m_2} (V_{ij}^{l_1 m_1 l_2 m_2})^2 \tag{21}$$

where the matrix $\{V_{ij}^{l_1 m_1 l_2 m_2}; i = 1, N; j = 1, N\}$ is defined by:

$$\mathbf{V}^{\mathbf{l_1 m_1 l_2 m_2}} = (\mathbf{U}_{\mathbf{A}}^{\mathbf{l_1 m_1}})^{\tau} \mathbf{U}_{\mathbf{B}}^{\mathbf{l_2 m_2}} \tag{22}$$

in which $\mathbf{U}_{\mathbf{A}}^{\mathbf{l_1 m_1}}$ and $\mathbf{U}_{\mathbf{B}}^{\mathbf{l_2 m_2}}$ are the eigenvectors of the atomic exchange matrices $\mathbf{Q}_{\mathbf{A}}^{\mathbf{l_1 m_1}}$ and $\mathbf{Q}_{\mathbf{B}}^{\mathbf{l_2 m_2}}$. If $\mathbf{V}^{\mathbf{l_1 m_1 l_2 m_2}}$ is the unit matrix the exchange energy is reduced to

the Coulomb energy and the exchange moments are equal to the electrostatic moments. As a result the eigenvector matrix $\mathbf{U}^{lm}$ characterizes the exchange nature of the interaction.

5.5 Conclusion

In this study we have focused on the exchange component of the Hartree–Fock energy or, more precisely, the "exchange-correlation" counterpart of the Kohn–Sham energy. In standard force fields this energy is absorbed in the parameterized bonded part, and not explicitly included in the non-bonded part. Our results show, however, that the mid-range intermolecular exchange energy is significant (tens of kJ mol^{-1}) and is, therefore, a crucial part of the total energy. We presented the use of a spherical tensor multipole expansion to compute exchange energy and force for a set of small molecules and dimers. The multipole moments defined by the overlap of molecular orbitals converge well for the exchange and in many cases better than the Coulomb energy. The exchange force, however, seems to converge more slowly. The exchange energy associated with $L = 1$ and multiplied by the internuclear distance is proportional to a previously established bond-order index.

References

1 W.D. Cornell, Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Jr., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. *J.Am.Chem.Soc.* **1995**, *117*, 5179–5197.

2 B.R. Brooks, Brucoleri R.E., Olafson B.D., Slater D.J., Swaminathan S., Karplus M. *J.Comp.Chem.* **1983**, *4*, 187–217.

3 W.R. Scott, Huenenberger P.H., Tironi I.G., Mark A.E., Billeter S.R., Fennen J., Torda A.E., Huber T., Krueger P., van Gunsteren W.F. *J.Phys.Chem.* **1999**, *103*, 3596–3607.

4 J.W. Ponder, Case D.A. *Adv. Protein Chem.* **2003**, *66*, 27–85.

5 R.F.W. Bader, *Atom in Molecules. A Quantum Theory*, Oxford Univ. Press (1990).

6 P.L.A. Popelier, *Atoms in Molecules. An Introduction*, Pearson, London, Great Britain (2000).

7 S. Brodersen, Wilke S., Leusen F.F.J. and G. Engel *Phys.Chem.Chem.Phys.* **2003**, *5*, 4923–4931.

8 S.L. Price, *Rev.in Comp.Chem.* (Editors: K.B. Lipkowitz and D.B. Boyd), Wiley, New York, USA, 2000.

9 S.L. Price, Harrison R.J., Guest M.F. *J. of Comp.Chem.* **1989**, *10*, 552–567.

10 A.J. Stone, Price S.L. *J.Phys.Chem.* **1988**, *92*, 3325–3335.

11 G.M. Day, Motherwell W.D.S., Jones W. *Crystal Growth & Design* **2005**, *5*, 1023–1033.

12 W.T.M. Mooij, Leusen F.J.J. *Phys.Chem.Chem.Phys.* **2001**, *3*, 5063–5066.

13 V. Marcon, Raos G. *J.Phys.Chem.B* **2004**, *108*, 18053–18064.

14 E.R. Batista, Xantheas S.S., Jonsson H. *J.Chem.Phys.* **1999**, *111*, 6011–6015.

15 C. Sagui, Pedersen L.G., Darden T.A. *J.Chem.Phys.* **2004**, *120*, 73–87.

16 J.P. Ritchie, Copenhaver A.S. *J.of Comp.Chem.* **1995**, *16*, 777–789.
17 U. Koch, Egert E. *J.Comp.Chem.* **1995**, *16*, 937–944.
18 W.A. Sokalski, Keller D.A., Ornstein R.L., Rein R. *J. of Comp.Chem.* **1993**, *14*, 970–976.
19 E.R. Batista *J.Chem.Phys.* **2000**, *112*, 3285–3292.
20 D.S. Kosov, Popelier P.L.A. *J. Phys.Chem.A* **2000**, *104*, 7339–7345.
21 D.S. Kosov, Popelier P.L.A. *J.Chem.Phys.* **2000**, *113*, 3969–3974.
22 P.L.A. Popelier, Rafat M. *Chem.Phys.Lett.* **2003**, *376*, 148–153.
23 M. Rafat, Popelier P.L.A. *J.Chem.Phys.* **2005**, *123*, 204103-1,7.
24 P.L.A. Popelier, Joubert L., Kosov D.S. *J.Phys.Chem.A* **2001**, *105*, 8254–8261.
25 L. Joubert, Popelier P.L.A. *Phys.Chem.Chem.Phys.* **2002**, *4*, 4353–4359.
26 P.L.A. Popelier, Kosov D.S. *J.Chem.Phys.* **2001**, *114*, 6539–6547.
27 R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press, 2nd edn, SanDiego (1992).
28 M. Rafat, Popelier P.L.A. *J.Chem.Phys.* **2006**, *124*, 144102-1,7.
29 D.A. Varshalovich, Moskalev A.N., Khersonskii V.K., *Theory of Angular Momentum*, World Scientific (1988).
30 A.J. Stone, *The Theory of Intermolecular Forces*, Clarendon, Oxford, Great Britain (1996).
31 C. Haettig *Chem.Phys.Lett.* **1996**, *260*, 341–351.
32 C. Haettig, Hess B.A. *Molec. Phys.* **1994**, *81*, 813–824.
33 A.M. Pendas, Francisco E., Blanco M.A. *J.Comp.Chem.* **2005**, *26*, 344–351.
34 X. Fradera, Austen M.A., Bader R.F.W. *J.Phys.Chem. A* **1999**, *103*, 304–314.
35 J. Poater, Sola M., Duran M., Fradera X. *Theor Chem Acc* **2002**, *107*, 362–371.
36 Y.G. Wang, Werstiuk N.H. *J.Comp.Chem.* **2002**, *24*, 379–385.
37 N.C. Handy, Schaefer III H.F. *J.Chem.Phys.* **1984**, *81*, 5031–5033.
38 A.D. Becke *J.Chem.Phys.* **1993**, *98*, 1372–1377.
39 W. Koch, Holthausen M.C., *A Chemist's Guide to Density Functional Theory*, Wiley–VCH, Weinheim, Germany (2000).
40 GAUSSIAN03 Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Montgomery, J.A.J.; Vreven, J.T.; Kudin, K.N.; Burant, J.C.; Millam, J.M.; Iyengar, S.S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J.E.; Hratchian, H.P.; Cross, J.B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Ayala, P.Y.; Morokuma, K.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Zakrzewski, V.G.; Dapprich, S.; Daniels, A.D.; Strain, M.C.; Farkas, O.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Ortiz, J.V.; Cui, Q.; Baboul, A.G.; Clifford, S.; Cioslowski, J.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Challacombe, M.; Gill, P.M.W.; Johnson, B.; Chen, W.; Wong, M.W.; Gonzalez, C.; Pople, J.A. In Gaussian, Inc., Pittsburgh PA, 2003.

6
The ELF Topological Analysis Contribution to Conceptual Chemistry and Phenomenological Models

Bernard Silvi and Ronald J. Gillespie

6.1
Introduction

In the opinion of the celebrated mathematician and philosopher of science René Thom, chemical concepts such as those of bonds and valence lack scientific content [1]. He pointed out that the vocabulary used by chemists to describe matter at a microscopic level was often ambiguous. In fact most concepts have their origin in models developed at the beginning of the twentieth century and earlier which are, therefore, not fully consistent with quantum mechanics. To reconcile the chemical description of matter with the postulates of quantum mechanics it is necessary to build a mathematical model. This mathematical model is not unique, however, because different spaces (geometrical direct space, momentum space, Hilbert space) and different mathematical theories external to quantum mechanics can be used for this purpose.

The choice of the geometrical space has been pioneered by Raymond Daudel with the loge theory [2–4] the applicability of which remained limited to very small systems because the underlying mathematical theory, Shannon's information theory [5], requires evaluation of the N-particle distribution function. The theory of dynamical systems [6, 7] is a powerful method of analysis of space which has been convincingly introduced in chemistry by Richard Bader with the quantum theory of atoms in molecules (QTAIM) [8]. This theory is based on analysis of the gradient field of the electron-density distribution which enables the partitioning of the molecular space into basins associated to each atom.

The purpose of this chapter is to provide a digest of the topological theory of chemical bonding with particular emphasis on epistemological aspects. QTAIM has made important contributions to conceptual chemistry, for example the definition of the atom within a molecule, of the bond critical point, of the bond path, and of the charge concentrations or depletions which emerge from analysis of the Laplacian of the charge density $L(\mathbf{r})$ and which make a link with the valence shell electron pair repulsion (VSEPR) model [9–15]. To provide evidence for basins which correspond to bonds and lone pairs, however, one must perform dynamic

system analysis of the gradient vector field of another function [16, 17], the electron localization function [18]. These methods collectively form what Malcolm and Popelier call quantum chemical topology (QCT) [19], which can be regarded as an important part of conceptual chemistry (by conceptual chemistry we mean the branch of chemistry concerned with analysis of the concepts used by chemists, with their definitions and underlying models). In this chapter we will introduce the electron localization function (ELF) in a very chemical fashion and show how its topology can be used to revisit different phenomenological models of bonding or of molecular geometry.

6.2 Why ELF and What is ELF?

The description provided by chemistry considers a molecule as an assembly of atoms linked by bonds. An atom in a molecule consists of a kernel (the nucleus and the inner shell electrons) and valence electrons gathered in the valence shell. The structure of the kernel and the possible numbers of electrons belonging to the valence shell are given by the position of the element in the periodic table. A molecule usually has fewer electrons than the sum of the populations of the valence shells of its atoms, because some of the valence electrons may be shared in two or more valence shells. Such electrons are said to be bonding electrons whereas the other valence electrons are nonbonding. The arrangement of the electrons in the valence shells constitutes the chemical electronic structure. In this description the bonding arises from shared electrons.

One of the objectives of Lewis's theory of valence [20, 21] is to predict the most probable structure with the help of additional rules such as the octet rule and the rule of two. Lewis's approach emphasizes the electron pair as a key concept. It is worth noting that an N-electron system has at most $N/2$ electron pairs in the chemical description and $N(N-1)/2$ in the quantum mechanical description. The chemical approach is not appropriate because many concepts lack a clear definition and because it has no mathematical model behind it.

To recover the Lewis picture within the QCT framework, it is assumed there is a local function, hereafter called the *localization function*, whose gradient field enables the partitioning of the molecular space into adjacent non-overlapping volumes hereafter called *basins* in the following way:

1. The space occupied by an atom (with $Z > 2$) is divided into an inner region, the *core basin*, encompassing the nucleus, and an external region, the atomic valence shell, gathering its *valence basins* which may extend to infinity.
2. A valence basin may be shared by the valence shells of several atoms.
3. There is a high probability of finding $Z - n_v$ electrons within a core basin where n_v is the ordinal number of the group of the periodic table to which the element belongs (in other words the conventional number of valence electrons).

4. There is a high probability of finding an even number of electrons in a valence basin belonging to a closed-shell system.

The number of electrons $\overline{N}(\Omega_A)$ in each basin Ω_A is given by the integrated density:

$$\overline{N}(\Omega_A) = \int_{\Omega_A} \rho(\mathbf{r})\, d\mathbf{r} \tag{1}$$

which is, alternatively, the expectation value of the basin population operator $\overline{N}(\Omega_A)$ [22], i.e.:

$$\overline{N}(\Omega_A) = \langle \Psi | \hat{N}(\Omega_A) | \Psi \rangle \tag{2}$$

$\overline{N}(\Omega_A)$ is, therefore, a statistical average, the variance of which provides a measure of the electron localization. Zero variance corresponds to perfect localization within Ω_A whereas a large variance is an indication of a poor localization. The variance localization criterion can be generalized to any volume and, in particular, to a sampling volume v around a reference point of coordinate $\mathbf{r}$. The variance $\sigma^2(\mathbf{r})$ is expressed as:

$$\sigma^2(\mathbf{r}) = \overline{N}_{\perp}(\mathbf{r}) + \overline{N}_{//}(\mathbf{r}) - \overline{N}^2(\mathbf{r}) + \overline{N}(\mathbf{r}) \tag{3}$$

in terms of the population $\overline{N}(\mathbf{r})$, and of the opposite spin and the same spin pair populations:

$$\overline{N}_{\perp}(\mathbf{r}) = \int_v (\pi^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_{21}) + \pi^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_{21}))\, d\mathbf{r}_1\, d\mathbf{r}_{21} \tag{4}$$

$$\overline{N}_{//}(\mathbf{r}) = \int_v (\pi^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_{21}) + \pi^{\beta\beta}(\mathbf{r}_1, \mathbf{r}_{21}))\, d\mathbf{r}_1\, d\mathbf{r}_{21} \tag{5}$$

respectively. To check the ability of $\sigma^2(\mathbf{r})$ to locate the basin boundaries, let us consider the very simple model of four electrons distributed in two identical cubic boxes, A and B, of volume V_0, sharing a face, we further impose the condition there is always an opposite spin pair in each box and that the electron density is constant within the boxes. In this circumstance the electron density and the pair functions have the expressions:

$$\rho(\mathbf{r}) = \frac{2.0}{V_0} = \rho \quad \text{for } \mathbf{r} \in \text{A or B} \tag{6}$$

$$\pi^{\sigma\sigma}(\mathbf{r}, \mathbf{r}') = \begin{cases} 0.0 & \text{for } \mathbf{r} \text{ or } \mathbf{r}' \text{ in the same box} \\ \dfrac{1.0}{V_0^2} & \text{for } \mathbf{r} \text{ and } \mathbf{r}' \text{ in different boxes} \end{cases} \tag{7}$$

$$\pi^{\sigma\sigma'}(\mathbf{r},\mathbf{r}') = \frac{1.0}{V_0^2} \quad \text{for } \mathbf{r} \text{ and } \mathbf{r}' \in \text{A or B} \tag{8}$$

which imply that both $\nabla\rho$ and $\nabla^2\rho$ are identically zero and that:

$$\overline{N}(\mathbf{r}) = \rho v \tag{9}$$

and

$$\overline{N}_{\perp}(\mathbf{r}) = \frac{1}{2}\rho^2 v^2 = \frac{1}{2}\overline{N}^2(\mathbf{r}) \tag{10}$$

and thus

$$\sigma^2(\mathbf{r}) = \overline{N}(\mathbf{r}) - \frac{1}{2}\overline{N}^2(\mathbf{r}) + \overline{N}_{//}(\mathbf{r}) \tag{11}$$

If v is chosen to constrain $\overline{N}(\mathbf{r})$ to be constant, the only varying quantity in $\sigma^2(\mathbf{r})$ is $\overline{N}_{//}(\mathbf{r})$, which is zero if the sampling volume is totally in one box and which is maximum when it is equally shared by the two boxes, that is when the reference point belongs to the boundary. The integrated same spin pair density over a sampling volume around a reference point is thus a convenient descriptor of pair formation in the sense of Lewis's model. The spin pair composition [23] defined as:

$$c_{\pi}(\mathbf{r}) = \overline{N}(\mathbf{r})^{-8/3}\overline{N}_{//}(\mathbf{r}) \tag{12}$$

is a quantity independent of $\overline{N}(\mathbf{r})$ for small $\overline{N}(\mathbf{r})$ values.

It has been shown that ELF [18] is an excellent approximation to this function when put in the Lorentzian form $\eta(\mathbf{r}) = (1 + c_p{}^2(\boldsymbol{r}))^{-1}$ which confines the values of $\eta(\mathbf{r})$ in the [1, 0] interval. ELF was originally conceived as a measure of the Fermi hole curvature [18, 24] of an Hartree–Fock wavefunction. It has the advantage that it can be expressed analytically in terms of basis functions in all practical cases where the wavefunction is expressed in terms of orbitals whereas the spin pair composition must be calculated numerically. ELF has been alternatively interpreted in terms of local excess kinetic energy because of Pauli repulsion [25], in terms of localized orbitals [26], and, recently, as the nonadditive (inter-orbital) Fisher information contained in the electron distribution [27].

6.3 Concepts from the ELF Topology

The ELF topology has been extensively used for study of chemical bonding [28–80], aromaticity problems [81–88], reactivity [89–91], and chemical reactions [92–105]. Several review papers have already been published [86, 106–108]. The ELF

theory provides very appealing pictures of bonding and a very "chemical" population analysis.

6.3.1 The Synaptic Order

Topological partitioning of the ELF gradient field [16, 17] yields basins of attractors which can be thought of as corresponding to atomic cores, bonds, and lone pairs. In a molecule one can find two types of basin:

1. core basins surrounding nuclei with atomic number $Z > 2$ and labeled C(A), where A is the atomic symbol of the element, and
2. the valence basins.

The valence basins are characterized by the number of atomic valence shells in which they participate. This number is called the synaptic order [109]. There are, therefore, monosynaptic, disynaptic, trisynaptic basins, etc. Monosynaptic basins, labeled V(A), correspond to the lone pairs of the Lewis model, and polysynaptic basins to the shared pairs of the Lewis model. In particular, disynaptic basins, labeled V(A, X), correspond to two-center bonds, trisynaptic basins, labeled V(A, X, Y), to three-center bonds, etc. The valence shell of a molecule is the union of its valence basins. Because hydrogen nuclei are located within the valence shells of at least two atoms they are located in the corresponding polysynaptic basin. For example, the valence basin accounting for a C–H bond labeled V(C, H) contains a proton and is called, for this reason, protonated disynaptic. The valence shell of an atom, say A, in a molecule is the union of the valence basins whose label lists contain the element symbol A. Figure 6.1 shows two examples of trisynaptic basins – the protonated trisynaptic basins of diborane V(B, B, H), which provides a picture close to the "protonated double bond" of Pitzer [110], and the pentacoordinated carbons of $Al_2C_2H_{10}$, which involves three V(C, H) disynaptic basins and one V(Al, Al, C) trisynaptic basin.

6.3.2 The Localization Domains

In the context of ELF analysis the concept of a domain is very important, because it enables definition of the chemical units within a system and characterization of

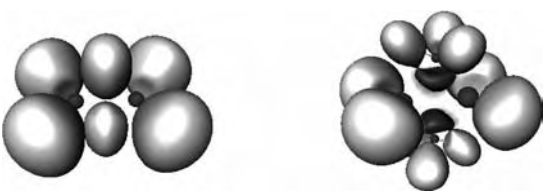


Fig. 6.1 Localization domains of B_2H_6 (left) and $Al_2C_2H_{10}$ (right).

the valence domains belonging to a given chemical unit. The mathematical properties of the gradient dynamical system do not, by themselves, provide the whole set of definitions necessary to describe the bonding in molecules. Other mathematically based approaches are therefore required for this purpose.

The topological concept of domain was introduced to chemistry by P. Mezey to recognize functional groups within organic molecules [111]. Generalized to ELF isovalues this concept has proven to be an efficient "generator" of clear definitions. Any subset of molecular space bounded by an external closed isosurface $\eta(\mathbf{r}) = f$ is a domain. An *f*-localization domain is such a subset, with the restriction that each point satisfies $\eta(\mathbf{r}) > f$. A localization domain which surrounds at least one attractor is called *irreducible*. If a localization domain contains more than one attractor it is reducible. An irreducible domain is a subset of a basin whereas a reducible domain is the union of subsets of different basins. Except for atoms and linear molecules, the irreducible domains are always filled volumes whereas the reducible domains can be either filled or hollowed volumes. If the value of $\eta(\mathbf{r})$ defining the bounding isosurface is increased a reducible domain splits into several domains each containing fewer attractors than the parent domain. Reduction of localization occurs at turning points, which are critical points of index 1 located on the separatrix of two basins involved in the parent domain. Ordering these turning points (localization nodes) by increasing $\eta(\mathbf{r})$ enables one to build tree-diagrams reflecting the hierarchy of the basins.

For any system there are low values of $\eta(\mathbf{r}) = f$ defining a unique composite parent domain. The successive reductions of localization will split this parent domain. Every branch which is a composite domain corresponds to one or more chemical species. A chemical unit is the union of the basins of the last appearing composite domain of a branch, if it is a filled volume. In a complex such as the weak hydrogen bonded system FH···ClH shown in Fig. 6.2 the first reduction yields two composite domains corresponding to the interacting moieties. Such a complex cannot be regarded as being *chemically bonded* [39]. In the same way, in an ionic pair such as NaCl, the first reduction yields domains corresponding to the cation and to the anion as shown in Fig. 6.3. In a molecule the initial parent domain first splits into core domains and a single valence domain which contains

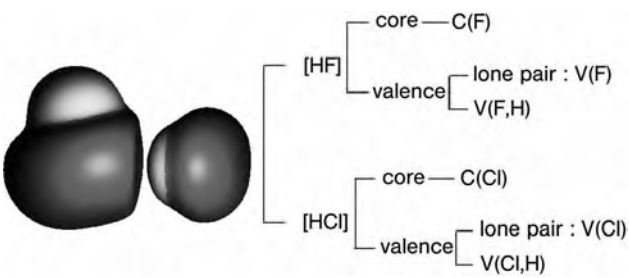


Fig. 6.2 $\eta(\mathbf{r}) = 0.06$ localization domains and reduction of the localization tree-diagram of FH···ClH.

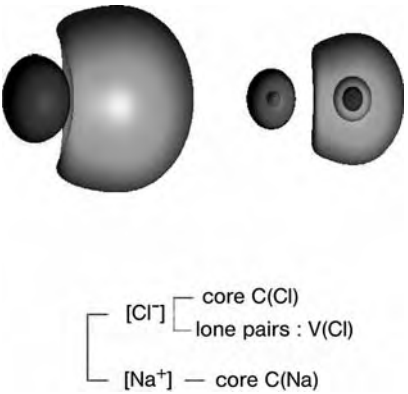


Fig. 6.3 $\eta(\boldsymbol{r}) = 0.05$ (left) and $\eta(\boldsymbol{r}) = 0.5$ (right) localization domains and reduction of the localization tree-diagram of NaCl.

all the valence attractors. The shape of this latter domain is that of a hollowed volume with as many holes as atomic cores in the molecule. Each hole contains a core domain as displayed for H_2CO in Fig. 6.4.

6.3.3
ELF Population Analysis

ELF population analysis relies on the integrated density over the basin volumes. The basin populations are the expectation values of the population operator introduced by Diner and Claverie [112]:

$$\hat{N}(\Omega_A) = \sum_i^N \hat{\gamma}(\mathbf{r}_i) \quad \text{with } \hat{\gamma}(\mathbf{r}_i) = \begin{cases} 1 & \mathbf{r}_i \in \Omega_A \\ 0 & \mathbf{r}_i \notin \Omega_A \end{cases} \hat{N}(\Omega_A) \tag{13}$$

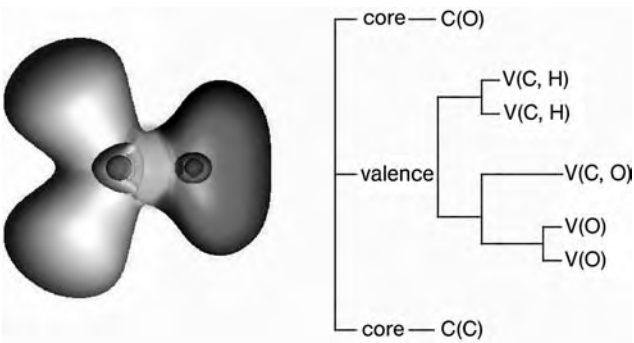


Fig. 6.4 $\eta(\boldsymbol{r}) = 0.4$ localization domains and reduction of the localization tree-diagram of H_2CO.

where N denotes the total number of electrons. The eigenvalues, $N(\Omega_A)$, of $\hat{N}(\Omega_A)$ belong to the series of integer values $0, \ldots, N$ and represent all the accessible numbers of electrons within Ω_A. The eigenvalues of the population operators of different basins are correlated, because they also obey the closure relationship:

$$\sum_A N(\Omega_A) = N \tag{14}$$

The expectation values of the population operators given by Eq. (2) can be expressed in terms of the volume integral of the one-electron probability distribution over the basins (Eq. 1). They are real numbers and can be understood as the average of measurements of the electron numbers $N(\Omega_A)$. They also fulfill a closure relationship, i.e.:

$$\sum_A \overline{N}(\Omega_A) = N \tag{15}$$

These eigenvalues and expectation values are, in fact, determined simultaneously. Each set of eigenvalues defines an accessible chemical electronic structure and the expectation values $\overline{N}(\Omega_A)$ can therefore be interpreted as weighted averages of mesomeric structures. The closure relation of the basin population operators enables one to perform a statistical analysis of the basin populations by definition of a covariance matrix [22]. The covariance operator is a matrix operator whose elements are deduced from the classical expression of the covariance:

$$\hat{\text{cov}}(\Omega_A, \Omega_B) = \hat{N}(\Omega_A)\hat{N}(\Omega_B) - \overline{N}(\Omega_A)\overline{N}(\Omega_B) \tag{16}$$

The covariance matrix element expectation values are the differences between the actual pair populations $\overline{\Pi}(\Omega_A, \Omega_B)$ and their "classical" analogs $\overline{N}(\Omega_A)\overline{N}(\Omega_B)$, or $\overline{N}(\Omega_A)(\overline{N}(\Omega_A) - 1)$ for the diagonal elements:

$$\langle \hat{\text{cov}}(\Omega_A, \Omega_A) \rangle = \overline{\Pi}(\Omega_A, \Omega_A) - \overline{N}(\Omega_A)(\overline{N}(\Omega_A) - 1) \tag{17}$$

$$\langle \hat{\text{cov}}(\Omega_A, \Omega_B) \rangle = \overline{\Pi}(\Omega_A, \Omega_B) - \overline{N}(\Omega_A)\overline{N}(\Omega_B) \tag{18}$$

The diagonal elements of the covariance matrix (the variances), are often denoted $\sigma^2(\overline{N})$, because they classically represent the square of the standard deviation σ.

For open-shell systems it is also very interesting to localize the unpaired electrons by calculating the integrated spin density over localization basins.

Although the topological representation enables rather satisfactory interpretation of the bonding, reliable descriptions in terms of the superposition of chemical structures are often very helpful, at least as explanatory models. As has been proposed in two previous papers [22, 81], the data provided by the topological

Table 6.1 Valence basin populations (e) and covariance matrix elements of H_2CO.

Basin	$\bar{N}(\Omega)$	$\langle \hat{cov}(\Omega, \Omega') \rangle$		
		V(C, H)	V(C, O)	V(O)
V(C, H)	2.13	0.52	−0.17	−0.07
V(C, O)	2.43	−0.17	1.16	−0.40
V(O)	2.52	−0.07	−0.40	0.98

analysis can be used to build such models and to evaluate their ability to describe the distribution of electrons. To illustrate this procedure we consider the formaldehyde molecule as an example. The valence basin populations and their covariances are reported in Table 6.1.

The three mesomeric structures considered are:

(a) $H_2C{=}O$
(b) $H_2C^+{-}O^-$
(c) $H_2C^-O^+$

The weights calculated for these structures are 0.24, 0.58, and 0.18 for (a), (b), and (c), respectively. They yield populations and covariance in good agreement with population analysis.

6.4 VSEPR Electron Domains and the Volume of ELF Basins

The VSEPR model relies on a distribution of the valence electron pairs among bonding and nonbonding electron domains [12, 13] which are defined as the regions of space in which the probability of finding an (opposite spin) electron pair or a large fraction of an electron pair is high. It is assumed that:

1. nonbonding domains or lone pairs are larger in size than bonding domains and are therefore more repelling;
2. the size of the bonding domains decreases when the electronegativity of the ligand increases and/or the electronegativity of the central atom decreases; and
3. multiple bonds have larger domains than a single bonds.

The topology of the Laplacian of the charge density has been invoked as a physical basis for the VSEPR model [113]. Qualitatively, the valence shell charge concentrations (VSCCs) of the central atom correspond to the electronic domains of the model [12, 13] and they have also been used to explain the geometries of non-

VSEPR molecules [114]. Malcolm and Popelier attempted to justify quantitatively the VSEPR assumptions on the relative domain sizes by considering the full topology of the Laplacian of the charge density [19]. They concluded that nonbonding domains are larger than bonding domains belonging to the same valence shell and that multiple bond domains are larger than single bond domains, but they found that the full topology of the Laplacian does not account for electronegativity effects.

There is, in principle, a one-to-one correspondence between the VSEPR electronic domains and the valence basins of the ELF function. To define finite volumes, the valence basins are limited by a bounding density isosurface. We have analyzed the volumes of the basins of approximately 150 molecules. Mathematically a basin is allowed to extend to infinity, which is not chemically meaningful; we have therefore limited molecular space by the density isosurface $\rho(\mathbf{r}) = 10^{-4}$ bohr^{-3}, which ensures that no more than 0.2*e* are omitted.

The volumes of the core basins of the main group elements of the four first periods range from 0.1 bohr^3 (Ne) to 149.5 bohr^3 (K). For the group 13–17, however, the largest value is only 16.5 (Ga). The core volumes of the fourth-period transition metals are in the range 50.5–20.5 bohr^3 (Sc and Ni respectively). It is worth noting that in transition-metal molecules the core external shell is split into several basins [66] the volumes of which are of the order of a few bohr^3. Figures 6.5 to 6.8 depict the volumes of the V(X) lone-pair basins and the V(X, H) and V(X, X) bonding basins as functions of the Allred and Rochow electronegativity of X [115]. They reveal very good correlations that are almost period-independent. For a given element the largest deviation occurs for the V(N) basin of CH_3CN whose population is 3.2*e*. As a general rule, basin volumes decrease when the

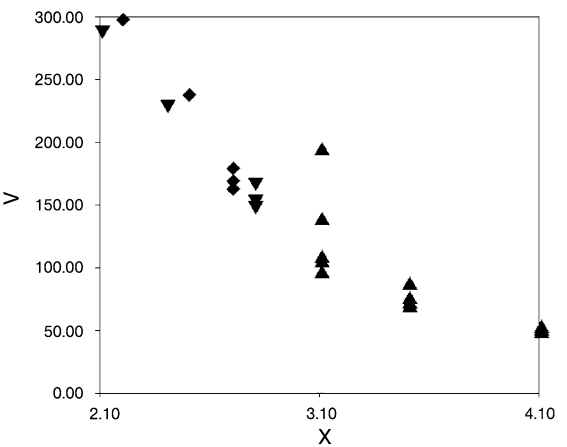


Fig. 6.5 Dependence of V(X) volume (bohr^3) on χ(X) for HF, HCl, HBr, H_2O, H_2S, H_2Se, NH_3, PH_3, AsH_3, CH_3F, CH_3Cl, CH_3Br, CH_3OH, CH_3NH_2, H_2CO, H_2CNH, F_2, Cl_2, and Br_2. ▲ second period elements, ▼ third period elements, ◆ fourth period elements.

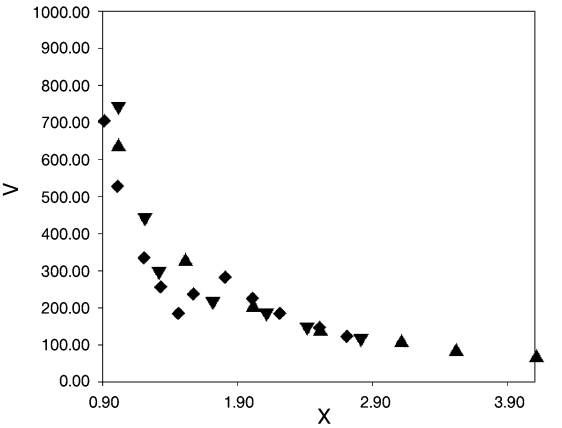


Fig. 6.6 Dependence of V(A, H) volume (bohr^3) on χ(A) in AH_n hydrides. ▲ second period elements, ▼ third period elements, ◆ fourth period elements.

electronegativity of X increases. The shrinking of the V(X) basin with electronegativity is a consequence of the radial decay of the electron density, which is governed by the core net charge. For the V(X, H) and V(C, X) bonding basins the observed behavior can be explained by the electronegativity differences which determine the ionic character of the bond. For $\chi(X) > \chi(H)$ or $\chi(X) > \chi(C)$ the ionic contributions involving either H^+ or CR_3^+ tend to reduce the basin volume whereas the anionic contribution H^- or CR_3^- act in the opposite direction when $\chi(X) < \chi(H)$ or $\chi(X) < \chi(C)$. Finally, the decrease of the V(X, X) basin volume

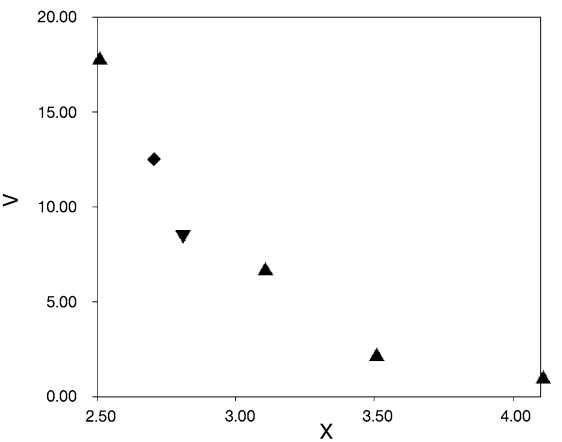


Fig. 6.7 Dependence of V(A, A) volume (bohr^3) on χ(A). ▲ second period elements, ▼ third period elements, ◆ fourth period elements.

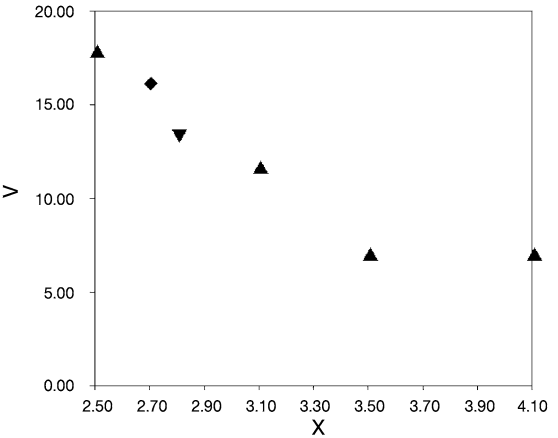


Fig. 6.8 Dependence of V(C, X) volume (bohr^3) on $\Delta\chi$. ▲ second period elements, ▼ third period elements, ◆ fourth period elements.

with increasing electronegativity is a consequence of the charge-shift nature of the bonds [73, 116].

To check the VSEPR assumptions we consider:

1. the V(X), V(X, H), and V(C, X) basins of CH_3XH_n (X = F, Cl, Br, O, N; $n = 0, 1, 2$); and
2. the V(X) and V(X, H) basins of H_2O, H_2S, H_2Se, NH_3, PH_3 and AsH_3.

The volumes of these basins are reported in Table 6.2. Around the X center in the CH_3XH_n series the V(C, X) basins are always smaller than the V(X) basins, in

Table 6.2 V(X), V(X, H), and V(C, X) basin volumes (a.u.).

Molecule	V(X)	V(X, H)	V(C, X)
NH_3	108.9	109.0	
PH_3	287.4	178.0	
AsH_3	297.4	183.8	
H_2O	75.6	86.0	
H_2S	228.2	139.5	
H_2Se	231.3	146.0	
CH_3NH_2	96.8	103.5	11.6
CH_3OH	69.1	80.6	7.0
CH_3F	48.5		7.0
CH_3Cl	152.1		13.3
CH_3Br	168.6		16.1

agreement with assumption 1, that lone-pair basins are always larger than the bonding domains on the same center. They are also in agreement with assumption 2, that the size of a bonding domain decreases when the electronegativity of the ligand increases. For the period-2 molecules NH_3, H_2O, CH_3NH_2, and CH_3OH, however, the V(X) basins are almost equal to, or slightly smaller than the V(X) basins, which is not in agreement with assumption 2.

This is also true for the third and fourth period hydrides but not for the period-2 hydrides, for which the V(X) basins are almost equal to, or slightly smaller than, the (X, H) basins, which is not in agreement with assumption 1. The ELF analysis therefore supports the first VSEPR assumption, i.e. that nonbonding or lone-pair domains are larger than bonding domains in the valence shell of the same central atom, but it is not in accord with the second assumption for hydrogen ligands. Gillespie and Robinson [117] have, however, recently concluded from a survey of bond angles and their relationship with ligand radii that the most important factor determining bond angles are ligand sizes rather than the size of bonding pair domains. They have shown, however, that the VSEPR and ligand–ligand repulsion models usually lead to the same conclusions, except for hydrogen ligands the particularly small size of which is the key factor in determining geometry. For a given ligand and central atoms of the same group, the volume of the V(A, X) basin increases with period number and, to some extent, with decreasing electronegativity. For example, the V(Si, F) basin volume is twice that of V(C, F).

6.5
Examples of the Correspondence Between ELF Basins and the Domains of the VSEPR Model

6.5.1
Octet Molecules

6.5.1.1 Hydrides (CH_4, NH_3, H_2O)

In these molecules ELF shows:

1. a basin corresponding to the central atom core; and
2. four, three, and two disynaptic basins, respectively, corresponding to the bonding domains of the VSEPR model, and one or two monosynaptic basins corresponding to the lone-pair domains of the VSEPR model.

Hydrides are unique in that hydrogen has no core. Accordingly there is no core basin. These basins could also be considered to correspond to protonated lone pairs, i.e. as protonated monosynaptic basins. As shown in Table 6.3, the populations of the A–H bonds follow the expected trend – they decrease as the electronegativity of the central atom increases, that is, when bond polarity increases.

Table 6.3 V(A), V(A, X), and V(X) basin populations (e), (A = C, N, O, Si; X = H, F, Cl).

Molecule	V(A)	(A, X)	V(X)
CH_4		1.98	
NH_3	2.16	1.91	
H_2O	2.27	1.66	
CF_4		1.31	6.54[a]
$SiCl_4$		1.61	6.40[a]

[a] The F and Cl lone pairs are merged in a single basin.

6.5.1.2 AX_4 (CH_4, CF_4, $SiCl_4$)

For these molecules ELF reveals a basin corresponding to the core, four disynaptic basins corresponding to the four bonding basins of the VSEPR model, and a monosynaptic basin on each ligand corresponding to the sum of the lone-pair domains of the VSEPR model. Because of the necessary cylindrical symmetry around a single bond the nonbonding electrons do not, in fact, form three nonbonding pairs but have a most probable location in a ring perpendicular to the bond axis and the basin has a corresponding toroidal shape. In CF_4 and $SiCl_4$ the population of each of the disynaptic basins is less than two electrons whereas the populations of the monosynaptic basins are correspondingly larger (Table 6.3). These populations are characteristic of polar bonds, the extent to which the population of the disynaptic basins is less than two electrons is a measure of their bond polarity.

6.5.1.3 AX_3E and AX_2E_2 (NCl_3, OCl_2)

The ELF localization domains of these molecules are displayed in Fig. 6.9. They look similar to the corresponding hydrides except that the ligands also have a monosynaptic basin as in the AX_4 molecules. It can happen that the expected disynaptic basin does not exist, presumably because of the large electronegativity difference. The monosynaptic (lone-pair) basins seem larger than the bonding basins, and this is verified by their populations, consistent with the VSEPR model

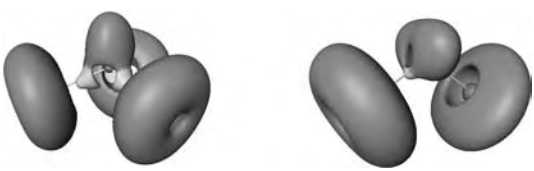


Fig. 6.9 ELF = 0.75 localization domains of NCl_3 (left) and OCl_2 (right).

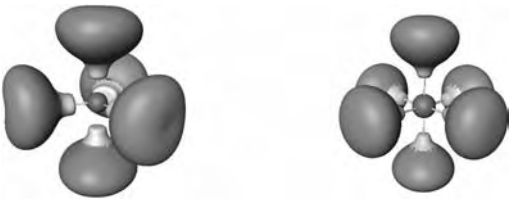


Fig. 6.10 ELF = 0.75 localization domains of PCl_5 (left) and SF_6 (right).

that postulated that the lone-pair domains take up more space around the core than the bonding domains.

6.5.2 Hypervalent Molecules

6.5.2.1 PCl_5 and SF_6

Hypervalent molecules are those in which the octet rule is violated because their Lewis structures have more than eight electrons in the valence shell of the central atom: PCl_5 and SF_6 are typical examples. ELF plots for these molecules are shown in Fig. 6.10. In each example there are core basins for each atom, disynaptic basins corresponding to the domains of the VSEPR model, and a monosynaptic basin on each ligand. The disynaptic basins have a population of less than two electrons, showing that the bonds are polar.

6.5.2.2 SF_4 and ClF_3

According to the VSEPR model these molecules have, respectively, one and two lone-pair domains in the valence shell of the central atom. The ELF plots in Fig. 6.11 show the monosynaptic basins corresponding to these lone-pair domains.

6.5.2.3 AX_7 and AX_6E Molecules

The structures of molecules with seven ligands cannot be predicted with certainty by the VSEPR model. The points-on-a-sphere model shows that the three most probable structures are the pentagonal pyramid, the mono-capped distorted octa-

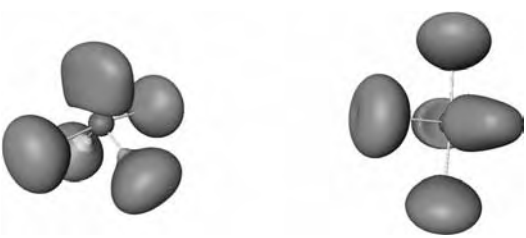


Fig. 6.11 ELF = 0.75 localization domains of SF_4 (left) and ClF_3 (right).

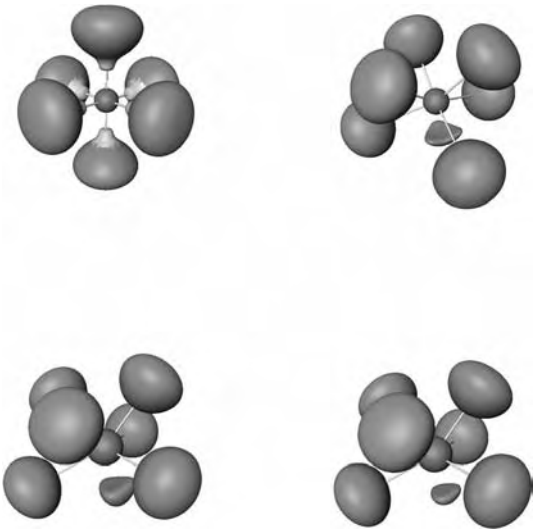


Fig. 6.12 ELF = 0.75 localization domains of BrF_6^- (top left), SeF_6^{2-} (top right), IF_6^- (bottom left), and XeF_6 (bottom right).

hedron, and the mono-capped trigonal prism. All known seven-coordinated molecules, for example IF_7 and TeF_7^-, have a pentagonal bipyramid structure. It might be expected, therefore, that AX_6E molecules, for example $TeCl_6^{2-}$, BrF_6^-, SeF_6^{2-}, and XeF_6, would have a structure a based on the pentagonal bipyramid with a lone pair in an axial position (the least crowded position), that is to say, a pentagonal pyramid. No AX_6E molecules have this shape, however; they are either octahedral (O_h, for example $TeCl_6^{2-}$ and BrF_6^-, or a have a C_{3v} distorted octahedral shape, for example SeF_6^{2-} and XeF_6. It has been proposed that in the octahedral molecules the ligands take up all the space in the valence shell leaving no room for nonbonding electrons, which form an outer shell of the core. In C_{3v} molecules the crowding of the ligands is not as great, so there is room for some nonbonding density in the valence shell forming what has been described as a ''weak lone pair''. The ELF plot for these molecules is in good agreement with this proposal. In Fig. 6.12 it can be seen there is a nonbonding basin in the valence shell of the central atom corresponding to the proposed weak lone pair. This basin is also involved in strong exchange with the core, consistent with the idea that the weak lone pair has partial lone-pair and partial core character.

6.5.3 Multiple Bonds

6.5.3.1 C_2H_4 and C_2H_2

ELF plots for these two molecules are given in Fig. 6.13. In C_2H_4 there are four disynaptic basins corresponding to the four C–H bonds, and a large disynaptic

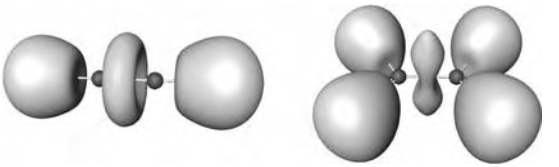


Fig. 6.13 ELF = 0.75 localization domains of C_2H_2 (left) and C_2H_4 (right).

basin corresponding to the C=C double bond. This basin has two maxima, one on each side of the CC axis, and splits into two basins at a sufficiently high value. These two basins are consistent with the VSEPR model of the double bond which consists of two bent bonds.

In the ELF plot for C_2H_2 there are two disynaptic basins corresponding to the two CH bond domains and a toroidal basin that surrounds the CC axis. This basin corresponds to the six bonding electrons. In the VSEPR model there would be three bonding pairs with their most probable locations located around the CC axis. As Linnett pointed out many years ago in his double quartet theory, however, there is no reason for the six bonding electrons to be localized as pairs, or indeed for any electrons that are not on the bond axis of a linear molecule to be fixed in localized pairs. Linnett proposed that the six bonding electrons have their most probable positions on a circle perpendicular to the bond axis with opposite spins alternating (Fig. 6.14). The observed toroidal basin in ELF corresponds to the Linnett model.

6.5.3.2 Si_2Me_4 and Si_2Me_2

According to the Lewis model these molecules have double and triple bonds just as in ethene and ethyne. There has, however, been much discussion about the real nature of these bonds, because their properties are not in agreement with these Lewis structures. In particular the molecules are not linear but have a *trans* bent shape. The ELF basins, however, give a clear picture of the electron distribution in these molecules, in which it can be seen that the formally multiple bond-

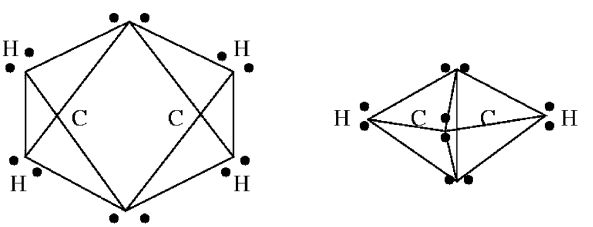
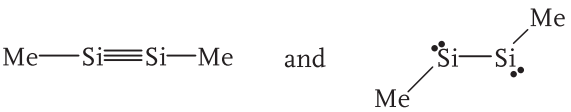


Fig. 6.14 Linnett representations of C_2H_4 (left) and C_2H_2 (right).

ing electrons also have some nonbonding or lone-pair character. The molecule Me_2Si_2 (Fig. 6.15) can, for example, be described as a resonance hybrid of:



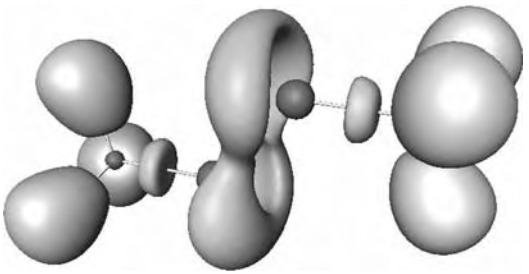


Fig. 6.15 ELF = 0.75 localization domains of Si_2Me_2.

This structure is consistent with the suggestion that only the most electronegative elements can hold two or three mutually repelling electron pairs close together in a multiple bond – these are principally the second period elements, C, N, and O, which do, indeed, form the vast majority of known multiple bonds.

6.6 Conclusions

ELF topological analysis provides a mathematical bridge between quantum mechanics and chemistry which relies on the one hand on statistical interpretation and on the other hand on the theory of dynamical systems. This approach shares the dynamical system theory, a common mathematical method of partitioning, with the quantum theory of atoms in molecules. The difference is in the nature of the potential function used and, therefore, the nature of the properties investigated. QTAIM is rightly claimed to be rooted in physics rather than in chemistry and the objective of its partition scheme is, accordingly, to define open quantum systems within which the virial theorem is valid. QTAIM does not, moreover, attempt to define the valence shell of an atom in a molecule. The goal of ELF partitioning is clearly to define basins of attractor corresponding to chemical concepts such as core and valence shells, bonds, and lone pairs. This is achieved at the expense of the physical meaning of the basins, which is lost. Another advantage of the ELF approach is that it provides pictures of the bonding in molecules which display pictorially the electron domains of the VSEPR model [13] and, in particular, provide evidence of the lone pairs [118]. ELF analysis has helped to refine

the conclusions of the VSEPR model, in particular for so-called hypervalent molecules [56], and to extend its applicability to the d^0 molecules of period 4 metals by considering basins of the core external shell [66]. Analysis of the basin volumes shows that the volumes of the core basins and of the structural features of their outermost shell are not always negligible. This analysis also enables rationalization of the subtle effects of electronegativity and, therefore, is a means of improving the VSEPR model.

References

1 R. Thom, *Paraboles et Catastrophes*, Flammarion, Paris, **1983**.
2 R. Daudel, *Les Fondements de la Chimie Théorique*, Gauthier–Villars, Paris, **1957**.
3 R. Daudel, in *Aspects de la Chimie quantique Contemporaine* (Eds. R. Daudel, A. Pullman), Editions du Center National de la Recherche Scientifique, Paris, **1971** p. 70.
4 C. Aslangul, R. Constanciel, R. Daudel, P. Kottis, in *Advances in Quantum Chemistry* (Ed. P.-O. Löwdin), Academic Press, New York, vol. 6, **1972** pp. 93–141.
5 C. E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, **1949**.
6 R. H. Abraham, C. D. Shaw, *Dynamics: The Geometry of Behavior*, Addison Wesley, **1992**.
7 R. H. Abraham, J. E. Marsden, *Foundations of Mechanics*, Addison Wesley, **1994**.
8 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford Univ. Press, Oxford, **1990**.
9 R. J. Gillespie, R. S. Nyholm, *Quart. Rev. Chem. Soc.* **1957**, *11*, 339–380.
10 R. J. Gillespie, *Molecular Geometry*, Van Nostrand Reinhold, London, **1972**.
11 R. J. Gillespie, I. Hargittai, *The VSEPR Model of Molecular Geometry*, Allyn and Bacon, Boston, **1991**.
12 R. J. Gillespie, *Chem. Soc. Rev.* **1991**, *21*, 59–69.
13 R. J. Gillespie, E. A. Robinson, *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 495–514.
14 R. J. Gillespie, *Coord. Chem. Chem. Rev.* **2000**, *197*, 51–69.
15 R. J. Gillespie, E. A. Robinson, *Chem. Soc. Rev.* **2005**, *34*, 396–407.
16 B. Silvi, A. Savin, *Nature* **1994**, *371*, 683–686.
17 U. Häussermann, S. Wengert, R. Nesper, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2069–2072.
18 A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397–5403.
19 N. O. J. Malcolm, P. L. A. Popelier, *Faraday Discuss.* **2003**, *124*, 353–363.
20 G. N. Lewis, *J. Am. Chem. Soc.* **1916**, *38*, 762–786.
21 G. N. Lewis, *Valence and the Structure of Atoms and Molecules*, Dover, New York, **1966**.
22 B. Silvi, *Phys. Chem. Chem. Phys.* **2004**, *6*, 256–260.
23 B. Silvi, *J. Phys. Chem. A* **2003**, *107*, 3081–3085.
24 J. F. Dobson, *J. Chem. Phys.* **1991**, *94*, 4328–4333.
25 A. Savin, O. Jepsen, J. Flad, O. K. Andersen, H. Preuss, H. G. von Schnering, *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 187–190.
26 J. K. Burdett, T. A. McCormick, *J. Phys. Chem. A* **1998**, *102*, 6366–6372.
27 R. Nalewajski, A. Koster, S. Escalante, *J. Phys. Chem. A* **2005**, *109*, 10038–10043.
28 A. Burkhardt, U. Wedig, H. G. von Schnering, A. Savin, *Z. anorg. allgem. Chem.* **1993**, *619*, 437–441.
29 A. Savin, B. Silvi, F. Colonna, *Can. J. Chem.* **1996**, *74*, 1088–1096.
30 M. E. Alikhani, Y. Bouteiller, B. Silvi, *J. Phys. Chem.* **1996**, *100*, 16092–16097.

31 B. Silvi, A. Savin, F. R. Wagner, in *Modelling of Minerals and Silicated Materials* (Eds. B. Silvi, P. D'Arco), Kluwer Academic Publishers, Dordrecht, vol. 15, topics in molecular organization and engineering ed., **1997** pp. 179–199.

32 S. Noury, F. Colonna, A. Savin, B. Silvi, *J. Mol. Struct.* **1998**, *450*, 59–68.

33 R. Llusar, A. Beltrán, J. Andrés, S. Noury, B. Silvi, *J. Comput. Chem.* **1999**, *20*, 1517–1526.

34 A. Beltrán, J. Andrés, S. Noury, B. Silvi, *J. Phys. Chem. A* **1999**, *103*, 3078–3088.

35 S. Berski, B. Silvi, Z. Latajka, J. Leszczynski, *J. Chem. Phys.* **1999**, *111*, 2542–2555.

36 D. B. Chesnut, A. Savin, *J. Am. Chem. Soc.* **1999**, *121*, 2335–2336.

37 I. Fourré, B. Silvi, P. Chaquin, A. Sevin, *J. Comput. Chem.* **1999**, *20*, 897–910.

38 G. Frison, A. Sevin, *J. Phys. Chem. A* **1999**, *103*, 10998–11003.

39 F. Fuster, B. Silvi, *Theoret. Chem. Acc.* **2000**, *104*, 13–21.

40 B. Silvi, C. Gatti, *J. Phys. Chem. A* **2000**, *104*, 947–953.

41 R. Choukroun, B. Donnadieu, J.-S. Zhao, P. Cassoux, C. Lepetit, B. Silvi, *Organometallics* **2000**, *19*, 1901–1911.

42 D. B. Chesnut, L. J. Bartolotti, *Chem. Phys.* **2000**, *253*, 1–11.

43 D. B. Chesnut, L. J. Bartolotti, *Chem. Phys.* **2000**, *257*, 175–181.

44 T. Kato, H. Gornitzka, A. Baceiredo, A. Savin, G. Bertrand, *J. Am. Chem. Soc.* **2000**, *122*, 998–999.

45 D. B. Chesnut, *J. Phys. Chem. A* **2000**, *104*, 7635–7638.

46 N. H. Werstiuk, H. M. Muchall, S. Noury, *J. Phys. Chem. A* **2000**, *104*, 11601–11605.

47 C. Fressigné, J. Maddaluno, A. Marquez, C. Giessner-Prettre, *J. Org. Chem.* **2001**, *65*, 8899–8907.

48 S. Berski, Z. Latajka, B. Silvi, J. Lundell, *J. Chem. Phys.* **2001**, *114*, 4349–4358.

49 S. Berski, B. Silvi, J. Lundell, S. Noury, Z. Latajka, in *New Trends in Quantum Systems in Chemistry and Physics* (Eds. J. Maruani, C. Minot, R. McWeeny, Y. G. Smeyers, S. Wilson), Kluwer Academic Publisher, Dordrecht, **2001** pp. 259–279.

50 J. Molina, J. Dobado, *Theor. Chem. Acc.* **2001**, *105*, 328–337.

51 D. B. Chesnut, *J. Comput. Chem.* **2001**, *14*, 1702–1711.

52 P. Fuentealba, A. Savin, *J. Phys. Chem. A* **2001**, *105*, 11531–115336.

53 H. Chevreau, I. Moreira, B. Silvi, F. Illas, *J. Phys. Chem. A* **2001**, *105*, 3570–3577.

54 D. B. Chesnut, L. J. Bartolotti, *Chem. Phys.* **2002**, *278*, 101–1102.

55 D. B. Chesnut, *Heteroatom Chem.* **2002**, *13*, 53–62.

56 S. Noury, B. Silvi, R. G. Gillespie, *Inorg. Chem.* **2002**, *41*, 2164–2172.

57 P. Mori-Sánchez, J. M. Recio, B. Silvi, C. Sousa, A. Martín Pendás, V. Luaña, F. Illas, *Phys. Rev. B* **2002**, *66*, 075103.

58 M. Feliz, R. Llusar, J. Andrés, S. Berski, B. Silvi, *New J. Chem.* **2002**, *26*, 844–850.

59 M. Kohout, F. R. Wagner, Y. Grin, *Theor. Chem. Acc.* **2002**, *108*, 150–156.

60 B. Silvi, F. Fuster, E. Kryachko, O. Tishchenko, M. T. Nguyen, *Molec. Phys.* **2002**, *100*, 1659–1675.

61 E. Chamorro, P. Fuentalba, A. Savin, *J. Comput. Chem.* **2003**, *24*, 496–504.

62 B. Silvi, J. Pilmé, F. Fuster, E. M. Alikhani, in *Metal–Ligand interactions in molecular, nano, micro, and macro-systems in complex environments, NATO-ASI series* (Eds. N. Russo, M. Witko), Kluwer, Dordrecht, **2003** pp. 241–284.

63 J. Pilmé, B. Silvi, M. Alikhani, *J. Phys. Chem. A* **2003**, *107*, 4506–4514.

64 H. Chevreau, C. Martinsky, A. Sevin, C. Minot, B. Silvi, *New J. Chem.* **2003**, *27*, 1049–1053.

65 J. R. B. Gomes, F. Illas, B. Silvi, *Chem. Phys. Lett.* **2004**, *388*, 132–138.

66 R. J. Gillespie, S. Noury, J. Pilmé, B. Silvi, *Inorg. Chem.* **2004**, *43*, 3248–3256.

67 H. Chevreau, *Chem. Phys. Lett.* **2004**, *400*, 59–61.

68 L. F. Pacios, *J. Phys. Chem. A* **2004**, *108*, 1177–1188.

69 T. M. Klapotke, B. Krumm, K. Polborn, I. Schwab, *J. Am. Chem. Soc.* **2004**, *126*, 14166–14175.

70 D. Chesnut, A. Crumbliss, *Chem. Phys.* **2005**, *315*, 53–58.

71 D. Chesnut, *Chem. Phys.* **2005**, *315*, 59–64.

72 F. Kraus, T. Hanauer, N. Korber, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7200–7204.

73 S. Shaiks, D. Danovich, B. Silvi, D. Lauvergnat, P. Hiberty, *Chem. Eur. J.* **2005**, *21*, 6358–6371.

74 J. Pilme, B. Silvi, M. E. Alikhani, *J. Phys. Chem. A* **2005**, *109*, 10028–10037.

75 B. Gomez, N. V. Likhanova, M. A. Dominguez Aguilar, O. Olivares, J. M. Hallen, J. M. Martinez-Magadan, *J. Phys. Chem. A* **2005**, *109*, 8950–8957.

76 D. Zubarev, A. Boldyrev, X. Li, L.-F. Cui, L.-S. Wang, *J. Phys. Chem. A* **2005**, *109*, 11385–11394.

77 W. Tiznado, O. B. Ona, V. E. Bazterra, M. C. Caputo, J. C. Facelli, M. B. Ferraro, P. Fuentealba, *J. Chem. Phys.* **2005**, *123*, 214302.

78 I. Vidal, S. Melchor, J. Dobado, *J. Phys. Chem. A* **2005**, *109*, 7500–7508.

79 J.-F. Boily, T. M. Seward, *Geochimica et Cosmochimica Acta* **2005**, *69*, 3773–3789.

80 J. Rosdahl, T. F. Fässler, L. Kloo, *Eur. J. Inorg. Chem.* **2005**, *14*, 2888–2894.

81 C. Lepetit, B. Silvi, Chauvin, *J. Phys. Chem. A* **2003**, *107*, 464–473.

82 C. Lepetit, V. Peyrou, R. Chauvin, *Phys. Chem. Chem. Phys.* **2004**, *6*, 303–309.

83 J. C. Santos, W. Tiznado, R. Contreras, P. Fuentealba, *J. Chem. Phys.* **2004**, *120*, 1670–1673.

84 J. Hernández-Trujillo, I. García-Cruz, J. M. Martínez-Magadán, *Chem. Phys.* **2005**, *308*, 181–192.

85 F. Kraus, N. Korber, *Chem. Eur. J.* **2005**, *11*, 5945–5959.

86 J. Poater, M. Duran, M. Solà, B. Silvi, *Chem. Rev.* **2005**, *105*, 3911–3947.

87 K. Knabel, T. M. Klapötke, H. Nöth, R. T. Paine, I. Schwab, *Eur. J. Inorg. Chem.* **2005**, *6*, 1099–1108.

88 J. Geier, H. Grützmacher, K. Exner, H. Prinzbach, *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 2433–2437.

89 F. Fuster, B. Silvi, *Chem. Phys.* **2000**, *252*, 279–287.

90 F. Fuster, A. Sevin, B. Silvi, *J. Phys. Chem. A* **2000**, *104*, 852–858.

91 F. Fuster, A. Sevin, B. Silvi, *J. Comput. Chem.* **2000**, *21*, 509–514.

92 X. Krokidis, S. Noury, B. Silvi, *J. Phys. Chem. A* **1997**, *101*, 7277–7282.

93 X. Krokidis, B. Silvi, M. E. Alikhani, *Chem. Phys. Lett.* **1998**, *292*, 35–45.

94 X. Krokidis, V. Goncalves, A. Savin, B. Silvi, *J. Phys. Chem. A* **1998**, *102*, 5065–5073.

95 X. Krokidis, B. Silvi, C. Dezarnaud-Dandine, A. Sevin, *New J. Chem.* **1998**, *22*, 1341–1350.

96 X. Krokidis, N. W. Moriarty, W. A. Lester, Jr., M. Frenklach, *Chem. Phys. Lett.* **1999**, *314*, 534–542.

97 X. Krokidis, R. Vuilleumier, D. Borgis, B. Silvi, *Molec. Phys.* **1999**, *96*, 265–273.

98 S. Berski, J. Andrés, B. Silvi, L. Domingo, *J. Phys. Chem. A* **2003**, *107*, 6014–6024.

99 M. C. Michelini, E. Sicilia, N. Russo, M. E. Alikhani, B. Silvi, *J. Phys. Chem. A* **2003**, *107*, 4862–4868.

100 M. C. Michelini, N. Russo, M. E. Alikhani, B. Silvi, *J. Comp. Chem.* **2004**, *25*, 1647–1655.

101 M. E. Alikhani, B. Silvi, *J. Mol.Struct.* **2004**, *706*, 3–6.

102 V. Polo, J. Andres, R. Castillo, S. Berski, B. Silvi, *Chem. Eur. J.* **2004**, *10*, 5165–5172.

103 V. Polo, J. Andrés, *J. Comp. Chem.* **2005**, *26*, 1427–1437.

104 J. C. Santos, V. Polo, J. Andrés, *Chem. Phys. Lett.* **2005**, *406*, 393–397.

105 J. C. Santos, J. Andrés, A. Aizman, P. Fuentealba, V. Polo, *J. Phys. Chem. A* **2005**, *109*, 3687–3693.

106 A. Savin, R. Nesper, S. Wengert, T. F. Fässler, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1809–1832.

107 B. Silvi, I. Fourré, E. Alikhani, *Monatshefte für Chemie* **2005**, *136*, 855–879.

108 M. E. Alikhani, F. Fuster, B. Silvi, *Structural Chemistry* **2005**, *16*, 203–210.
109 B. Silvi, *J. Mol. Struct.* **2002**, *614*, 3–10.
110 K. S. Pitzer, *J. Am. Chem. Soc.* **1946**, *67*, 1126–1132.
111 P. G. Mezey, *Can. J. Chem.* **1993**, *72*, 928–935.
112 S. Diner, P. Claverie, in *Localization and Delocalization in Quantum Chemistry* (Eds. O. Chalvet, R. Daudel, S. Diner, J. P. Malrieu), Reidel, Dordrecht, vol. II, **1976** pp. 395–448.
113 R. F. W. Bader, R. J. Gillespie, P. J. MacDougall, *J. Am. Chem. Soc.* **1988**, *110*, 7329–7336.
114 R. J. Gillespie, I. Bytheway, T.-H. Tang, R. F. W. Bader, *Inorg. Chem.* **1996**, *35*, 3954–3963.
115 A. L. Allred, E. G. Rochow, *J. Inorg. Nucl. Chem.* **1958**, *5*, 264.
116 S. Shaik, P. Maitre, G. Sini, P. C. Hiberty, *J. Am. Chem. Soc.* **1992**, *114*, 7861–7866.
117 R. J. Gillespie, E. A. Robinson, *Inorg. Chem.* **2004**, *43*, 2318.
118 D. B. Chesnut, *J. Phys. Chem. A* **2000**, *104*, 11644–11650.

Part II
Solid State and Surfaces

7
Solid State Applications of QTAIM and the Source Function – Molecular Crystals, Surfaces, Host–Guest Systems and Molecular Complexes

Carlo Gatti

7.1
Introduction

This chapter deals with the application of QTAIM to the solid state, except for the last section, in which the source function – a recently developed tool for the QTAIM study of the chemical bond from a somewhat original viewpoint – is introduced.

The chapter starts with an illustration of TOPOND software, which implements QTAIM for systems periodic in 0 to 3 dimensions, which covers polymers, surfaces, and crystals, besides molecules. The interface of TOPOND to the multipolar package XD is also mentioned, because it enables QTAIM analysis of experimentally derived electron densities. The chapter continues with an example of a didactic application of TOPOND to a study of crystal field effects on bonding and on molecular properties in the urea molecular crystal. Clean and chemisorbed semiconductor surfaces then serve as an example of the wealth of information provided by QTAIM about the effect of surface formation and reconstruction on the bonding and atomic properties of first surface layers. Guest–host systems are discussed as a last example, with emphasis on guest to/from host electron transfer and on the peculiar features of guest–host bonding interactions. The relevance of these key issues to materials science applications is briefly touched upon for thermoelectric materials.

In the last section, the source function (SF) is introduced and examples of its preliminary and future potential applications and extensions are presented. This function enables the value of the density at any point within a system to be equated to a sum of atomic contributions, thus enabling the properties of the density to be viewed from a totally new perspective. Although depending on the whole set of interactions within a system, a bond path is *topologically* associated only with the two atoms it connects. In contrast, the source function details how *all the other atoms* in a system, in addition to the two linked atoms, contribute to the accumulation of electron density along a bond path and, in particular, to BCP. It thus discloses nonlocal information on bonding and on complex bonding pat-

terns, analogously to the QTAIM delocalization index or the synaptic order of an ELF valence basin. One advantage of the SF over these two powerful interpretive tools is that it is directly amenable to experimental determination, since to evaluate it only the knowledge of the system's electron density and Laplacian is required.

7.2 QTAIM Applied to Solids – the TOPOND Package

TOPOND [1–3] has some specific, important features which are summarized below and which mark it out from other QTAIM software for the condensed phase (a list, not exhaustive, is given elsewhere [4, 5]). As an obvious prerequisite, application of QTAIM to the solid state implies a knowledge of at least the electron density and its derivatives in a representative portion of the system. If the system is periodic in nature this representative portion corresponds to the unit cell or, simply, to the unique part of it, the asymmetric unit. The usual practical implementations of QTAIM to solids usually work on a user's defined volume, which includes the basin of the unique atoms within the system and the basins of their bonded atoms. Precise information about the periodic nature of the system is usually lost this way, and application of QTAIM to periodic systems is essentially brought back to a topological study of the electron density of a large cluster of atoms extracted from the crystal. The electron density is either calculated analytically or simply given on a grid [4], whereas the ρ derivatives, in particular those of order greater than two, are only estimated numerically by the large majority of software. (Topological analysis of $\rho(\mathbf{r})$ and of $\nabla^2\rho(\mathbf{r})$ requires $\rho(\mathbf{r})$ derivatives up to the second and up to the fourth order, respectively.) A full implementation of QTAIM [6] also necessitates a knowledge of the one-electron density matrix (ODM) and of the main diagonal of the two-electron density matrix, the pair density; neither of these is usually available in QTAIM software for solid systems, however.

Most of the listed shortcomings are, on the contrary, simply absent from TOPOND-98 [1, 7], because of its intimate interface with the libraries of, and output from, the CRYSTAL-98 package [8]. (TOPOND-98 is currently interfaced with CRYSTAL-98. Linking to CRYSTAL-2003, the most recent release of CRYSTAL, is currently in progress and planned to be complete by the end of 2006.) The CRYSTAL software performs ab-initio calculations of the ground-state energy, electronic wave function, and properties of periodic systems in 0 (molecules), 1 (polymers), 2 (slabs), and 3 dimensions (crystals). Systems with different periodicity are treated on an equal footing in CRYSTAL, with the single particle wave functions being expanded, for any periodicity, as a linear combination of Bloch functions, each of which is defined in terms of local atomic orbitals. Space symmetry is fully exploited, with 230 space groups, 80 layer groups, 99 rod groups, and 45 point groups available to the user. These unique features of CRYSTAL automatically make TOPOND a powerful tool for application of QTAIM to

molecules, polymers, surfaces, and crystals, using a single software product. TOPOND works on electron densities obtained with similar accuracy for different aggregations of matter, and topologically analyzes these densities with the same kinds of algorithm and precision. For example, one may very easily assess how the topological and atomic properties of a molecule are modified when it becomes surrounded by equivalent molecules in a given crystalline arrangement, or how the bonding patterns and atomic properties of a bulk atom in a solid change when this atom is placed on, and forms the surface of, a solid.

The intimate interface with CRYSTAL enables TOPOND to take advantage of the full periodic geometrical machinery of the former software. For example, this implies that when a critical point (CP) is located in a crystal, a full list of neighboring atoms, with their exact cell locations and with their coordinates given in fractional or Cartesian form, is immediately available. Or that all atoms of equivalent symmetry are easily recognized and unnecessary calculations are avoided as much as possible.

The electron density $\rho(\mathbf{r})$, [2, 9]:

$$\rho(\mathbf{r}) = \sum_{\mathbf{g},\mathbf{l}} \sum_{\mu,\nu} \mathbf{P}^{\mathbf{g}-\mathbf{l}}_{\mu,\nu} \chi^{\mathbf{g}}_{\mu}(\mathbf{r}) \chi^{\mathbf{l}}_{\nu}(\mathbf{r}) \tag{1}$$

and, if needed, all its derivatives up to the fourth order are computed analytically by TOPOND. In Eq. (1) χ_μ and χ_ν are atomic orbitals, $\mathbf{g}$ and $\mathbf{l}$ are lattice vectors, $\mathbf{P}^{\mathrm{g}-1}$ is the ODM associated to orbitals located in crystal cells having relative position $\mathbf{g}-\mathbf{l}$, and $\chi^{\mathbf{g}}_{\mu}$ is an atomic orbital located in cell $\mathbf{g}$, but with same shape and the same fractional coordinates as the χ_μ orbital in the reference zero cell ($\mathbf{g}=0$). Evaluation of $\rho(\mathbf{r})$ is not performed using a single threshold distance criterion from the point r to select whether a $\chi^{\mathbf{g}}_{\mu}(\mathbf{r})$ orbital contribution is to be included or not into the quadruple sum yielding $\rho(\mathbf{r})$. Instead, different distance thresholds are used for each orbital, based on the magnitude of their value at $\mathbf{r}$ [9]. This means that the "cluster of atoms" built around r for evaluating $\rho(\mathbf{r})$ does not have a predetermined fixed size as in all other QTAIM implementations for solids, but a different size for each μ, ν orbital, reflecting its specific diffuseness. ODMs are available within TOPOND and quantities such as the kinetic energy densities K($\mathbf{r}$) or G($\mathbf{r}$), the virial density $V(\mathbf{r})$, or the electron localization function (ELF) [10] may be all evaluated exactly, without resorting to their approximate expressions [11] in terms of $\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$, and $\nabla^2\rho(\mathbf{r})$. Unfortunately, the Fermi hole and the ensuing localization/delocalization indices [12] are not yet computed by TOPOND, but their evaluation will, hopefully, be included in a future release of the software, at least for HF or Kohn–Sham type wavefunctions.

The TOPOND package presently comprises five sections [1, 3]. The first two implement topological analysis of $\rho(\mathbf{r})$ and of $\nabla^2\rho(\mathbf{r})$, respectively. The fourth section of TOPOND deals with evaluation of atomic basin boundaries and associated atomic basin properties and the last section performs grid evaluation of several scalar fields, including the ELF, and traces out molecular/crystal graphs or, generically, $\nabla\rho$ trajectories in selected molecular/surface/crystal planes. The third sec-

tion, which is currently being implemented and not yet released is concerned with evaluation of IAS properties. These include, among others, the integrated surface charge, the net flux of the total electric field, yielding $q(\Omega)$ via the Gauss theorem [13], and the external source function contribution to the density within an atomic basin [14]. Two different CP search algorithms are available, the conventional Newton–Raphson (NR) technique and an eigenvector following (EF) method, proposed by Popelier [15]. The EF method can be seen as an NR method with a suitable and locally defined shift for the NR step. It is thus much less sensitive than the NR method to the choice of good starting search points. The EF method can seek CPs of a given kind, irrespective of the structure of the Hessian of the scalar field at the starting search point – a feature that makes it particularly helpful for topological study of $\nabla^2\rho$, because this function varies quite rapidly. Separate searches for the different kinds $(3,-3;3,-1;3,+1;3,+3)$ of CPs are implemented in TOPOND for both the ρ and the $\nabla^2\rho$ fields. A fully automatic search strategy, able to find sequentially all kinds of electron density CPs, is also available. This strategy makes use of the relevant EF step prescription for each kind of CP searched for in sequence. A CP search on a grid, defined in the asymmetric unit, may be also exploited when the Morse topological relationship [16], given by Eq. (2):

$$n - b + r - c = 0 \tag{2}$$

(where n, b, r and c are the total number of nuclear, bond, ring and cage CPs) is not fulfilled by the set of CPs found using the fully automatic search. The CP search on a grid is usually found to be 2 to 3 orders of magnitude more demanding computationally than the automatic search [3]. If one adopts suitable grid sizes, however, it seldom misses CPs, even when very flat density distributions occur, as in the metals. Atomic interaction lines (ρ field) and atomic graphs ($\nabla^2\rho$ field) are determined by TOPOND by tracing the associated steepest ascent/descent $\nabla\rho$ or $\nabla(\nabla^2\rho)$ paths using a fifth-order Runge–Kutta method with monitoring of local truncation error and adaptive step-size control. Correct parameterization of the algorithm (desired accuracy at each step, initial step size, oscillation control close to the 3, −3 attractors) enables tracing of the correct atomic interaction lines for metals also; for these the occurrence of nonnuclear attractors (NNAs) [17] is more the rule than the exception and the network of interaction lines is, consequently, rather complicated and largely unstable with regard to changes in the computational model or in the cell parameters [1, 3, 18, 19].

7.2.1 QTAIM Applied to Experimental Densities: TOPXD and XD Packages

During the last decade, QTAIM has increasingly been applied to crystalline systems [20, 21]. This is, on the one hand, because of the technical developments that have made X-ray diffraction a unique tool for mapping the charge density in

crystals and, on the other hand, because of noticeable improvements in ab-initio periodic approaches [22, 23], which have enabled calculation of reliable electron densities even for crystals having a large number of atoms (>50–100) in the unit cell. One may safely say that QTAIM is the primary standard theory used by the X-ray density community to discuss bonding in crystals [21]. This would not have been possible if software packages implementing the QTAIM for experimental densities had not been developed, documented, and made generally available in the past decade. One such software package is the TOPXD program [24, 25], which enables complete topological analysis of experimental charge densities on the basis of the Hansen–Coppens multipole formalism [26]. It is based on TOPOND-98, but with subroutines for geometrical calculation and density evaluations rewritten in the XD package convention [27, 28]. XD is the most widely distributed package for experimental charge density multipole refinement and TOPXD is fully integrated in the most recent version, XD 5.01, [27, 28]. The main features of TOPXD are those of TOPOND, with more extended documentation, friendly input style, increased speed in evaluating the IAS, and added facilities for their 3D visualization. The experimental electron density and its derivatives up to order 2 are calculated analytically, whereas derivatives of third and fourth order are obtained with great precision as a numerical finite-difference approximation of the first and second-order analytical derivatives [24]. In contrast with TOPOND, properties at a given point **r** are calculated by including density contributions from "pseudoatoms" which lie within a given distance threshold of **r**. This may limit the accuracy of the calculated properties and requires a check that the properties being computed converge relative to the distance threshold [27]. The ODM is not available within TOPXD and XD and all properties depending on this matrix (Section 7.2) can clearly not be computed.

One of the most important reasons for the popularity of QTAIM is that a large part of this theory uses, operationally, only information contained in the electron density $\rho(\mathbf{r})$, which enables unbiased comparison of theoretical and static experimental densities, irrespective of their diverse origin and of the different approximations one makes to derive them [20, 21, 29]. Such a comparison may provide information about the quality of experimental data and the suitability of the multipolar model used to project the reciprocal space representation of ρ to its real-space counterpart [30, 31]. Conversely, it may reveal deficiencies of the theoretical approach [30–32], for example poor treatment of the electron correlation, the use of an insufficiently flexible basis set, or the adoption of a pseudopotential with core–valence separation which is too crude.

The performance of the adopted multipolar model may be tested as follows [24, 31, 33]. First, a multipole refinement, keeping the positional parameters fixed and thermal parameters set to 1, is performed using structure factors F generated with an ab-initio periodic calculation to simulate the X-ray diffraction data set. Topological analysis is then performed on the density from the multipole model and the results compared with those obtained using the primary theoretical density. It has been shown that the multipole model may significantly bias the topolog-

ical results, because of the limited flexibility of the radial functions used in the multipolar analysis. Indeed, the observed discrepancies between theoretical and experimental, X-ray-derived topological properties are usually found to decrease substantially when the theoretical densities are projected into the multipole density functions by refinement of the theoretical F. All of this analysis may be performed easily by using CRYSTAL, TOPOND, and TOPXD/XD software in combination.

7.3 QTAIM Applied to Molecular Crystals

Electron distributions of crystals are an amazing source of information about the weaker and the less conventional atomic interactions [21], besides that provided about standard chemical bonds. Typical of molecular crystals is the simultaneous occurrence of normally strong intramolecular bonds, and of generally weak intermolecular contacts, with properties of both kinds of interaction being mutually affected, in contrast with the situation for the isolated molecule or the simple case of gas-phase molecular aggregates (dimers, trimers, ...). QTAIM has proved to be a very powerful tool for isolating, detailing the weak intermolecular interactions responsible of molecular crystal formation and for quantitative characterization of the effect of these interactions on intramolecular bonding.

Using the urea crystal as an example, answers will be provided to several simple questions:

1. How important are packing effects on intramolecular bonds?
2. Does the packing have different impact on the different atoms/chemical groups present in the molecule?
3. How large is the enhancement of the molecular dipole on crystallization?
4. How can each oxygen atom in the urea crystal be involved in four N–H···O hydrogen bonds (HBs)?
5. How does the global molecular volume contraction observed in the solid result from the individual atomic volume change on crystallization?

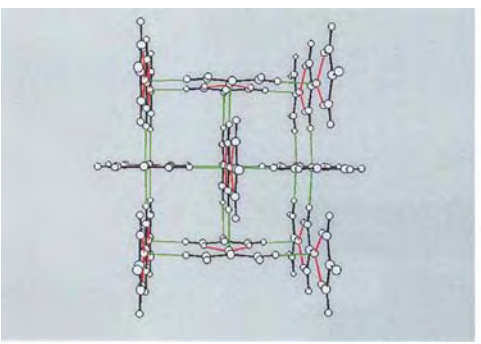
Before illustrating the situation for urea, it is worth mentioning that much more complex QTAIM applications to molecular crystals have appeared. To save space, however, these studies cannot be discussed here. I merely quote Refs [34] and [35] that address the interesting problem of the *nature* and *function* of the weak CH···O intermolecular interactions in crystals by analyzing the experimental and theoretical densities of the 3,4-bis(dimethylamino)-3-cyclobutene-1,2-dione (DMACB) crystal. This system is characterized by 23 unique intermolecular and intramolecular CH···O interactions and by no other type of stronger, and thus successfully competing, HB. References [34] and [35] discuss basic questions such as:

1. Does the existence or the absence of an CH···O CP reflect specific geometrical features of a CH···O contact in the DMACB crystal?
2. Can the bonded CH···O contacts in this crystal be classified as true HB?
3. Do the weak intermolecular HBs induce a large molecular dipole moment enhancement on crystallization, as typically found in molecular crystals tied by the much stronger NH···O and OH···O bonds?
4. Do the crystal and procrystal densities differ in the topological features of their CH···O contacts and can the CH···O bond energies thus be reliably retrieved from the BCP properties alone?

The reader is referred to the original papers to discover why three "YES" and one "NO" were the answers to these questions.

7.3.1
Urea

In the crystal structure (space group $P\bar{4}2_1m$), the urea molecules are linked to each other through HBs to form infinite planar tapes (O–H″ length 2.06 Å, HBs shown in red in Fig. 7.1) [36]. Adjacent tapes are mutually orthogonal and oriented in opposite directions, their cohesion being provided by another set of HBs (O–H′ length 1.99 Å, shown in green in Fig. 7.1). Each oxygen atom is in-



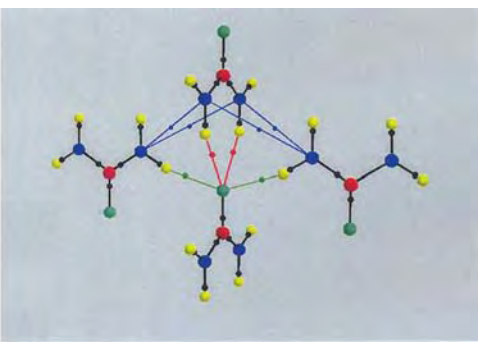


Fig. 7.1 The urea crystal. Left: The view is approximately along the *c* axis. Molecules are linked to each other through HBs (red lines) to form infinite planar tapes. Adjacent tapes are mutually orthogonal, oriented in opposite directions and tied together by another set of HBs (green lines). Right: Intramolecular and intermolecular bond paths (BCPs: small dots) for an urea molecule linked through HBs with another molecule along a tape and with two molecules of two neighboring orthogonal tapes. Bond paths in blue are associated with the shorter N–N contacts (see text). Each oxygen atom is involved in four HBs, two within the planar tape and two with neighboring orthogonal tapes.

volved in four HBs, two within the tape, and two with neighboring tapes, giving rise to a relatively high experimental sublimation energy (21 ± 0.5 kcal mol^{-1}). The urea structure seems to provide the only instance of a carbonyl atom which accepts four N–H···O hydrogen bonds [36]. In addition to the expected intramolecular and intermolecular interactions, application of QTAIM to the experimental geometry crystal periodic wavefunction (RHF/6-31G**) reveals [3] two unique long N···N intermolecular interactions between orthogonal (N–N, 3.433 Å, shown in blue in Fig. 7.1, right) and parallel (N–N, 4.267 Å) molecular tapes. The complete set of CPs in the unit cell (16*n*, 34*b*, 26*r*, and 8*c*) fulfils Morse's relationship [16] and it agrees with a recent experimental determination using synchrotron diffraction data and detailed multipolar analysis [37]. (The CP set found in the first QTAIM study of the urea crystal [2] also fulfilled Morse's relationship, but did not include the BCPs and RCPs related to the shorter of the two unique N···N contacts. With the adoption of automated CP searches in TOPOND-98 (Section 7.2), the complete set of CPs reported in the text was found [3, 21]. This set has been carefully checked with the grid search option of TOPOND-98.) The CP set is obtained by assigning to each unique CP found by TOPOND-98 a multiplicity equal to the multiplicity of the Wyckoff position where this unique CP is located.

7.3.1.1 **Urea: Packing Effects**

We start by obtaining an idea of the size of packing effects on molecular electron density (ED) compared with the density changes resulting from molecule formation. Figure 7.2 shows ρ contour plots and $\nabla\rho$ trajectory maps in the molecular plane of a urea molecule in the crystal, for three different ED models. From top to bottom we see how the ED changes on going from the independent atom model (IAM) density [26], which is given by the sum of atomic densities with atoms being placed at the same positions as in the crystal, to the density resulting from superimposition of the EDs of isolated molecules placed at crystal positions [8], and, finally, to the "true" crystal ED. The ED changes because of the *interaction density* – the crystal density minus the superimposition of the density of isolated molecules – are, indeed, barely detectable whereas those resulting from the *deformation density* – the crystal density minus the IAM model – are much more evident. This is, for example, apparent from the noticeable change in the size of the atomic basins. Despite its smallness, the interaction density is responsible for the large (37%, vide infra) enhancement of the molecular dipole magnitude $|\mu|$ in the crystal [2] and comparison of the atomic basins of the middle and bottom panels of Fig. 7.2 (right) immediately reveals how sensitive the values of $|\mu|$ are to the location of the atomic boundaries and to the atomic ED distribution changes. The IAS determinations and the subsequent atomic basin integrations must be very accurate to enable proper evaluation of the molecular dipole enhancement $|\Delta\boldsymbol{\mu}|$ occurring in the crystal. The changes of some topological properties at BCPs on passing from the IAM model to the isolated molecule at crystalline geometry (CG), then to the noninteracting molecules model, and finally to the crystal density are listed in Table 7.1. When the topological properties are analyzed,

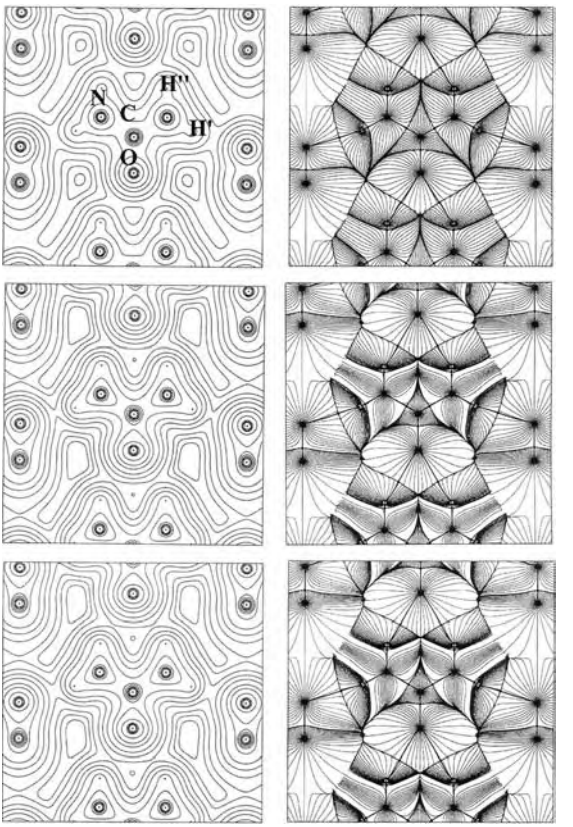


Fig. 7.2 Electron density contour plots (left) and $\nabla\rho$ trajectories maps (right) in the molecular plane of a urea molecule in the crystal for three different ED models. Top: IAM model; middle: noninteracting molecules model; bottom: crystal density periodic RHF model. In all models the molecules are placed at positions in the crystal and have crystal geometry. The noninteracting molecules and the crystal periodic RHF densities look very much alike despite the 37% enhancement in the crystal of the molecular dipole magnitude. In the $\nabla\rho$ trajectory maps, note the NH″···O bond paths linking molecules along a tape (central part of each map) and the NH′···O bond paths providing cohesion of this tape with the two neighboring orthogonal tapes (left and right of each map).

the effects resulting from the interaction density become clearly detectable. They are roughly *one order of magnitude lower* than those resulting from the deformation density for both intramolecular and intermolecular bonds. The molecular density and the sum of the noninteracting molecular densities are *indistinguishable* for intramolecular bonds. Conversely, the HBs, which are obviously absent in the isolated molecule, have, in the noninteracting molecules model, BCP properties already close to those in the crystal. This shows that the effect of packing on the intramolecular is more evident than that on the intermolecular ED distribution

Table 7.1 Changes of BCPs properties in urea with change of model density.[a,b]

Bond X–Y	ρ_b	$(\nabla^2\rho)_b$	$(\lambda_3)_b$	ε_b
C–O	0.299	0.17	1.02	0.03
	0.392	−0.55	1.49	0.07
	0.392	−0.55	1.49	0.08
	0.381	−0.33	1.75	0.00
C–N	0.264	−0.20	0.52	0.04
	0.341	−0.94	0.73	0.05
	0.341	−0.94	0.72	0.05
	0.349	−1.15	0.53	0.10
N–H′, 1.009 Å	0.235	−0.39	1.05	0.01
	0.345	−1.92	0.78	0.06
	0.345	−1.91	0.77	0.06
	0.344	−1.95	0.88	0.05
N–H″, 1.005 Å	0.237	−0.40	1.07	0.01
	0.351	−1.92	0.71	0.06
	0.351	−1.92	0.71	0.06
	0.349	−1.97	0.87	0.05
O···H′, 1.992 Å	0.029	0.08	0.15	0.06
	0.022	0.07	0.12	0.05
	0.022	0.07	0.12	0.07
O···H″, 2.058 Å	0.026	0.07	0.13	0.05
	0.019	0.07	0.11	0.01
	0.019	0.07	0.11	0.04

[a] All quantities except distances in a.u. For each bond the first entry refers to the IAM model, the second to the molecular density, the third to the superimposition of noninteracting molecular densities, and the fourth to the crystal density. The urea molecule is at crystal geometry ([36], 12 K geometry) in all models. The hydrogen bonds have only three entries because they are clearly absent for the isolated molecule.
[b] RHF/6-31G** wavefunction when appropriate.

[34]. The ED for the IAM model and for noninteracting molecules in the crystal were evaluated by use of the PATO and MOLSPLIT options, respectively, of CRYSTAL-98. The associated ODMs were then input to the TOPOND-98 software.

The percentage changes in the BCP properties of urea upon crystallization relative to those for the isolated molecule with the crystal geometry are listed in Table 7.2. (A full discussion of BCP and atomic property changes on crystallization, relative to the gas-phase molecule both with the crystal geometry and with the gas-phase optimized geometry (with C_{2v} constraint), is given elsewhere [2]) Easy chemical understanding of these changes is immediately obtained by using the well known Bader and Essen [38] classification of bonds into shared-shell ($\nabla^2\rho_b < 0$) and closed-shell ($\nabla^2\rho_b > 0$) interactions, according to the sign of

Table 7.2 Percentage changes in the BCP properties and changes in the atomic electron population and out-of-plane atomic quadrupole moment of urea on crystallization.[a,b,c]

X–Y	ΔR_x%	$\Delta\rho_b$%	$\Delta(\nabla^2\rho)_b$%	$\Delta(\lambda_3)_b$%	$\Delta\varepsilon_b$%	Ω	ΔN(Ω)	$\Delta Q_\perp(\Omega)$
C–O	−1.0	−2.8	−40.0	+17.4	−95.9	C	+0.032	−0.07
C–N	+2.0	+2.3	+22.3	−27.4	+110.6	O	+0.092	+0.08
N–H′	+1.5	−0.3	+1.6	+12.8	−20.3	N	+0.087	+0.36
N–H″	+2.9	−0.6	+2.6	+22.5	−25.0	H′	−0.053	
						H″	−0.100	

[a] Refs. [2, 21]; data refer to urea at the 12K neutron crystal geometry [36].
[b] For a given property P, ΔP% is evaluated as $\Delta P\% = [(P_{crystal} - P_{molecule})/P_{molecule}] \times 100$ whereas ΔP (P = N, $Q_\perp$) is given by $\Delta P = P_{crystal} - P_{molecule}$.
[c] R_x, the BCP distance from the X atom, is positive (negative) when the BCP is displaced towards Y (X) in the crystal compared with the molecule in crystal geometry. $Q_\perp(\Omega)$ is the out-of-molecular plane atomic quadrupole moment and a negative value denotes out-of-plane preferential accumulation of charge. A positive value for $\Delta Q_\perp(\Omega)$ indicates that in the crystal the electronic charge is partially moved from the out-of-plane to the molecular plane region.

$\nabla^2\rho_b$. Although too rigid, this dichotomous classification is usually valid for bonds among second-row atoms and it is particularly useful when applied on a relative basis [21]. It provides a set of quantitative indices [38] variation of which among a series of chemically related compounds or, as in this example, following a change of phase, pinpoint the effect these perturbations have on the nature of a given bond [21]. For example, if, on passing from the gas to the crystal phase the G_b/ρ_b and $(\lambda_3)_b$ values of a bond decrease and the ρ_b, $-\nabla^2\rho_b$, and $|\lambda_{1,2;b}|/(\lambda_3)_b$ values all increase, this bond, upon crystallization, has become more covalent; the reverse is true if the opposite changes occur. As shown in Table 7.2, for the C–N bond the ρ_b and $-\nabla^2\rho_b$ values increase and its parallel curvature largely decreases whereas the opposite changes occur for the C–O bond on passing from the molecule to the bulk. The N–H bonds undergo a negligible ρ_b decrease accompanied by a noticeable $(\lambda_3)_b$ increase. All these changes indicate that the C–O and N–H bonds become more polar and weaker whereas the covalence and strength of the C–N bond are enhanced on HBs formation in the crystal. Because of their link with the second derivative of the electron density, $(\nabla^2\rho)_b$, $(\lambda_3)_b$, and ε_b seem more sensitive indices of crystal field effects than ρ_b values, the changes of the second derivative-related quantities being at least an order of magnitude greater than those of ρ_b. For bonds such as C–O and C–N the greater their bond ellipticity, ε_b, the larger the extent of their π character. On crystallization, the double-bond character of C–N is substantially increased whereas for the C–O

bond this character is diminished, in agreement with the changes of the other BCP indices. As expected, the N quadrupole moment component perpendicular to the molecular plane is found to increase in the bulk ($\Delta Q_{\perp}(N) = 0.36$ au, Table 7.2), because electronic charge is removed from the out-of-plane lone pairs and placed in the C–N bond region [2]. (A negative value of $Q_{\perp}(\Omega)$ denotes out-of-plane preferential accumulation of charge [6].) The small increase found for $Q_{\perp}(O)$ in the bulk is a result of more uniform electron charge concentration in the nonbonding regions, because of formation of its four HBs (vide infra).

The BCP displacements upon change of phase (ΔR_x values in Table 7.2) are smaller for the more polar bonds, because the crystal field may less easily polarize this kind of bond. BCPs displace in such a way as to reduce the size of the H atoms and thus make them more positively charged in the bulk (Table 7.2). All heavy atoms, in contrast, have their population increased, the net result being a more polarized molecule, a larger dipole moment, and a flux in the crystal of 0.066 electrons from each amino group hydrogen donor to the carbonyl group acceptor. Table 7.3 shows that the molecular dipole moment magnitude $|\boldsymbol{\mu}|$ increases in the crystal by 37% and by 53% relative to the isolated molecules in the crystal and optimized geometry, respectively. As detailed in Table 7.3, the large dipole moment increase in the bulk is primarily as a result of a large increase of the magnitude of the charge-transfer component μ_{CT}, which contributes 88% and to 73% of the reported enhancements. The concomitant decrease of the first moment component, μ_A, because of the general reduction of atomic polarization on crystallization, serves only to slightly enhance the effect resulting from charge transfer (CT), because, as typically found, the atomic polarization contribution opposes the CT contribution [6].

Table 7.3 Urea. Changes of the molecular dipole moment and of its atomic and charge-transfer components on crystallization.[a,b]

Contribution	OG Molecule	CG Molecule	Crystal
μ_A, ($\Delta|\mu_A|$%)	0.71	0.54	0.45 (−16.7)
μ_{CT}, ($\Delta|\mu_{CT}|$%)	−2.52	−2.56	−3.22 (+25.8)
μ, ($\Delta|\mu|$%)	−1.81	−2.02	−2.77 (+37.1)

[a] Ref. [2]; OG and CG refer to gas-phase optimized geometry (C_{2v} constraint) and crystal geometry isolated molecules. All quantities in a.u.
[b] The dipole component parallel to the C_2 axis (directed from C to O atoms) is reported. Because of the C_{2v} symmetry of the molecule, this is the only nonvanishing component; μ_A and μ_{CT} are the first moment and the charge-transfer components of the total dipole moment μ.
The percentage dipole changes are evaluated as:
$\Delta|P|\% = [(|P_{crystal}| - |P_{CGmolecule}|)/|P_{CGmolecule}|] \times 100$, for
$P = \mu_A, \mu_{CT}, \mu$.

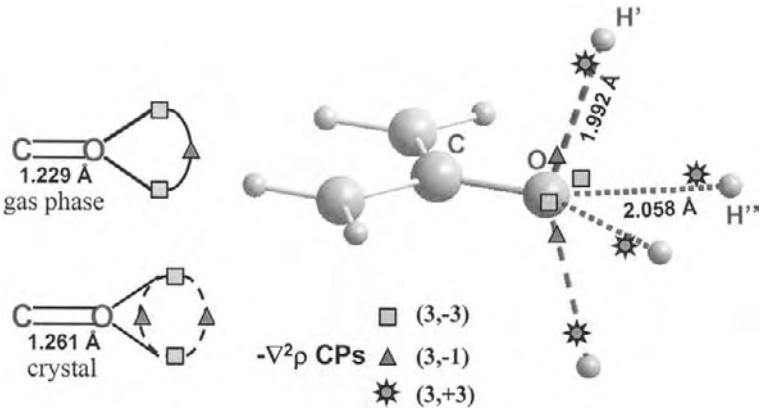


Fig. 7.3 Hydrogen bonds in the urea crystal and the $-\nabla^2\rho$ distribution. Left: On crystallization, the CO bond lengthens and the in-plane single saddle point linking the two O lone pair nonbonded $-\nabla^2\rho$ maxima bifurcates into two saddle points lying one above and the other below the molecular plane. Right: The O atom in the bulk may thus become involved in four NH···O hydrogen bonds (HBs). The in-plane HBs (dotted lines) are formed by aligning the $(3,+3)$ charge depletions (CDs) of the two H″ atoms of another molecule along a tape with the two O lone pair charge concentrations. The two out-of-plane HBs (dashed lines) exploit the alignment of the $(3,+3)$ charge depletions of H′ atoms of two neighboring orthogonal tapes with the two $(3,-1)$ O saddle points formed by the CO bond lengthening in the bulk. The saddle points are much better aligned with the H′ CDs than are the nonbonded maxima with the H″ CDs. This explains why the out-of-plane HBs are shorter and stronger than the in-plane HBs.

Careful analysis of the Laplacian CPs, in the valence shell charge concentration (VSCC) of the oxygen atom, provides an explanation of how this atom can form four NH···O bonds in the bulk [2]. In the urea molecule, in its C_{2v} optimized geometry, the two $(3,-3)-\nabla^2\rho$ nonbonded maxima (NBMs) associated with the O atom lone pairs are linked by a $(3,-1)-\nabla^2\rho$ saddle point, lying in the molecular plane (Fig. 7.3). The CO bond noticeably elongates in the crystal, from 1.229 to 1.261 Å, and as a result, this in-plane single saddle point bifurcates into two saddle points lying above and below the molecular plane (Fig. 7.3). Correspondingly, the two NBMs decrease in value, from 5.921 in the OG molecule to 5.669 a.u. in the crystal, whereas the charge concentration at the saddle point increases from 3.869 to 3.913 a.u., even though this point bifurcates in the bulk. The oxygen NBMs are the most electron-rich region in urea and their $-\nabla^2\rho$ value greatly exceeds the value of other charge concentrations (CCs) in the molecule, including those of the potentially competing NBMs (2.022 a.u.) associated with the N lone pair [2]. Conversely, the two $(3,-1)$ saddles interconnecting the oxygen NBMs are the second most electron-rich regions in the molecule and they are seen as maxima by the $(3,+3)$ charge depletions (CDs) of H atoms approaching the oxygen VSCC in a plane containing the C–O axis and perpendicular to the molecular plane. HBs can thus be formed either approaching the $(3,-3)$ NBMs or the two

saddles interconnecting them. Both options are exploited in bulk urea and each O atom thus becomes involved in four HBs. The oxygen VSCC changes in such a way as to form a torus of nearly uniform charge concentration in its nonbonded region, which increases its capability to form more than two HBs, and which agrees with the discussed increase of $Q_{\perp}(O)$ and the decrease of the C–O bond π character occurring in the bulk. It is worth noting that the bifurcation of the saddle point interconnecting the two oxygen NBMs also occurs in the isolated urea molecule, if the crystal geometry is adopted. Thus, the lengthening of the C–O bond in bulk urea, with the ensuing loss of the C–O bond π character and changes in the oxygen nonbonded regions, seems to be a key step in the onset of the observed 3D HB network in the bulk. A rationale for the out-of-plane HBs being shorter (1.992 Å) than the in-plane HBs is given elsewhere [2]. Basically, the oxygen $(3, -1)$ saddle points may be much more easily aligned to their facing hydrogen $(3, +3)$ CDs than is possible for the oxygen NBMs.

Table 7.4 lists the individual atomic volumes and their sum over a molecule or over the atoms of the –CO and $-NH_2$ functional groups for the urea molecule in the crystal and in the gas phase (CG and OG geometries). This table reveals how the atomic size of the different atoms or groups of atoms in the molecule changes on crystallization and how the molecular volume contraction observed in the solid occurs.

For the gas-phase molecules, the volumes determined using the 0.001 (V_1) and the 0.002 a.u. (V_2) density envelope are reported, and the basin total volumes, V_T, also are given for the crystal, because of their finite size in the condensed phase.

Table 7.4 Urea. Changes of the molecular, atomic, and functional group volumes on crystallization.[a]

Ω	OG Molecule		CG Molecule		Crystal		
	V_1	V_2	V_1	V_2	V_1	V_2	V_T
C	18.7	17.4	19.9	18.5	21.0	19.1	21.2
O	129.7	108.9	132.2	111.1	115.6	104.6	124.3
N	118.0	100.8	117.9	101.4	120.2	107.5	135.0
H′	24.7	18.9	24.6	18.8	16.8	15.3	17.7
H″	26.7	20.2	27.4	20.8	18.2	16.3	18.7
CO	148.4	126.3	152.1	129.6	136.6	123.7	145.5
NH_2	169.4	139.8	170.0	140.9	155.2	139.1	171.4
Molecule	487.3	406.0	492.2	411.5	447.0	401.8	488.2

[a] Ref. [2]; V_1 and V_2 are the volumes of the portion of the atomic basins where ρ is equal to or exceeds 0.001 or 0.002 a.u., respectively. V_T is the total volume of the atomic basin; its value is finite only in the bulk.

The computed total molecular volume in the crystal (488.2 a.u.) reproduces extremely well the volume per molecule in the unit cell (489.5). As a reaction of the molecules to the intermolecular exchange forces, the molecular volume reduces on crystallization by approximately 10% if the V_1 values are compared and by about 2% if the V_2 values are considered, with V_T being very close to V_1 in the gas phase. V_2 is approximately 90% and 83% of V_1 in the bulk and gas phases, respectively, indicating that the molecular density dies off more rapidly in the bulk than in the gas phase. Interestingly, the reported trends do not apply to the individual atomic basins, but they do to the –CO and $-NH_2$ groups. This suggests that the two functional groups *respond as the whole molecule* to the change of phase, despite their constituting atoms reacting in quite different ways to the same perturbation. The contraction of the molecular volume on crystallization originates primarily from contraction of the atomic basins directly involved in the hydrogen bonds. For other atoms a small expansion is observed. The substantial decrease of oxygen ($\Delta V_2 = -6\%$) and hydrogen (H′, $\Delta V_2 = -19\%$; H″, $\Delta V_2 = -28\%$) atomic volumes is related to the mutual penetration of their van der Waals envelopes after HB formation. Penetration is measured by the difference, ΔR, between the nonbonded radius of the H or of the O in the isolated molecule and the distance from the corresponding atoms to the HB CP in the crystal. (The nonbonded radius is taken as the average distance from the nucleus to the 0.001 a.u. contour in the isolated molecule [6].) The ΔR values are, as expected, higher for the shorter HB (H: 0.70; O: 0.56 a.u.) than for the longer HB (H: 0.59 and O: 0.54 a.u.) and the HB bond length difference results mostly in H′ being penetrated much more efficiently than H″.

7.4
QTAIM Applied to Surfaces

Several experimental techniques are now available for obtaining qualitative information on the geometrical reconstructions which occur at semiconductor surfaces. The electronic structure and, especially, accurate EDD of the surface are much less accessible experimentally [39, 40]. In contrast, not only do ab-initio periodical calculations of model surfaces usually reproduce the observed geometrical distortions satisfactorily, but also do afford a fairly accurate description of both the electronic structure and the EDD at the surface [41]. QTAIM analysis of the ab-initio wavefunctions then results in valuable, quantitative insight into the interplay between geometrical reconstructions, electronic relaxations, and bonding [42, 43].

QTAIM enables one to address very basic and important issues, for example:

- How is the bonding of the surface atoms affected?
- What is a ''dangling'' bond and how are its properties modified by surface reconstruction?
- How rapidly do the surface perturbations and the changes in these perturbations that result from the adsorption process

decay into the crystal in terms of the properties of the atoms in each of the succeeding layers?
- What are the charges on the surface atoms on a free surface and in the chemisorbed states?
- How much electronic charge is transferred and in what direction?
- What is the atomic origin and nature of the surface "double layer"?
- What atoms are exposed on a particular chemisorbed or clean surface?
- How can the surface character of an atom be defined?

Answers to these questions are given below for the Si (111)(1×1) clean and hydrogen-covered surfaces [42] and for the well known (2×1) reconstruction of the (111)(1×1) surface [43].

7.4.1
Si(111)(1×1) Clean and Hydrogen-covered Surfaces

The Si(111)(1×1)–H surface, in which all the silicon *dangling bonds* (DBs) are saturated by a single monolayer of hydrogen atoms, is one of the most ideal and simple semiconductor surfaces [44]. The clean Si(111)(1×1) surface, in which the DBs of the threefold-coordinated atoms of the first layer are not allowed to form bonds, is an *ideal* surface, which is not found in nature and which serves here and in the next discussed example for comparison purposes. These two surfaces were both described [42, 43] using the slab model (SLAB keyword) in the CRYSTAL program. A slab is a 2D periodic structure with two infinite periodic surfaces parallel to a crystalline plane (111, in our case), and a finite thickness, defined by the number of layers of atoms considered. Definition of a proper slab thickness is not a trivial task and the reader is referred to the original paper for technical details [42]. Stick-and-ball representations of the adopted centrosymmetric ten-layer slab models for the two surfaces are shown in Fig. 7.4 (top). In Tables 7.5 and 7.6 their bond and atomic properties are listed as a function of depth from surface and compared with corresponding data for the Si bulk and the SiH_4 molecule. Labeling of the Si atoms, with Si1 being an Si atom in the outermost Si layer, is reported in Fig. 7.4 (top). Bonding patterns in a plane normal to the slabs' surfaces and containing nuclei Si1 to Si5 are made evident by the $\nabla\rho$ trajectories portrayed in Fig. 7.4 (middle).

The large CT from Si1 to H induces noticeable asymmetry (Δ_{BCP} values, Table 7.5) of the surface Si1–Si2 back bonds which become stronger (ρ_b, $\nabla^2\rho_b$, λ_3 values, Table 7.5) than in the bulk, whereas in the clean slab these same bonds are weakened compared with the Si crystal, in accordance with the well-known propensity of the (111)(1×1) surface to evolve in favor of other reconstructions (Section 7.4.2). Analysis of the dependence of Si–Si bond properties on depth from the surface indicates that starting from Si2–Si3 bond in the hydrogen-

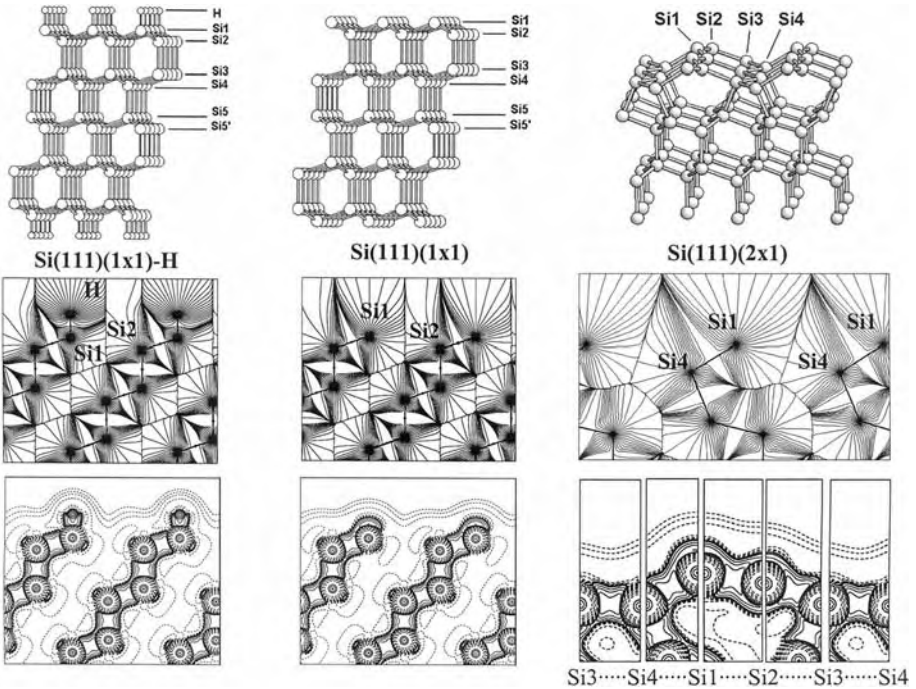


Fig. 7.4 Si(111)(1×1)–H, Si(111)(1×1) and Si(111)(2×1) surfaces. Top: Adopted slab models and labeling of the atomic layers. For the reconstructed Si(111)(2×1) surface, only half of the centrosymmetric slab is shown. On reconstruction, the surface dangling bonds are on nearest-neighbor sites rather than on next-nearest-neighbor sites. In Si(111)(2×1) there are two layers of chain-bonded atoms, a top layer (Si1 and Si2) of threefold-coordinated atoms and a lower layer (Si3 and Si4) of fourfold-coordinated atoms. Middle: $\nabla\rho$ trajectories and bonding pattern in planes normal to the slabs' surfaces and containing the Si1 to Si5 nuclei [Si(111)(1×1); Si(111)(1×1)–H] or (right) Si1–Si4 bonds [Si(111)(2×1)]. Bottom: $-\nabla^2\rho$ contour plots in the same planes of the $\nabla\rho$ trajectories for the Si(111)(1×1)–H and Si(111)(1×1) slabs. In the reconstructed Si(111)(2×1) surface (right), contour plots in five different planes normal to the surface are juxtaposed to show the whole bonding network among the atoms of the two outermost layers. Adapted, with permission, from Figs. 2 and 3 of Ref. [42], F. Cargnoni, C. Gatti, E. May, D. Narducci, *J. Chem. Phys.* 2000, *112*, 887–899 (Copyright 2000, American Institute of Physics) and from Figs. 1, 3, and 4 of Ref. [43], F. Cargnoni, C. Gatti, *Theor. Chem. Acc.* 2001, *105*, 309–322 (Copyright 2001, Springer).

covered and from Si3–Si4 bond in the clean surface, the Si–Si bonds become bulk-like.

The negatively charged H layer [q(H) = −0.736] in the hydrogen-covered slab is almost entirely balanced [q(H+Si1+Si2) = +0.015] by the first two Si layers, with more than 94% of the compensating charge coming from Si1 atoms [q(Si1) = +0.698]. Distortions in the charge of surface cells are often treated [45] in terms of a uniform macroscopic surface density of dipoles. In this representa-

Table 7.5 Relationship between bond properties and depth from surface for hydrogen-covered and clean Si(111)(1×1) slabs compared with silicon bulk and SiH_4.[a,b]

Bond	R_e (Å)	Δ_{BCP}	$\rho_b \times 100$	$\nabla^2\rho_b \times 100$	$(\lambda_3)_b \times 100$	ε
Si(111)(1×1)–H; *Si(111)(1×1) clean*						
Si–H	1.487	4.1	11.40	38.2	76.6	0.0
Si1–Si2	2.362	−1.4	8.83	−13.3	2.1	0.02
	2.362	*0.1*	*8.35*	*−11.2*	*2.7*	*0.05*
Si2–Si3	2.378	0.8	8.53	−12.1	2.2	0.0
	2.381	*−0.1*	*8.40*	*−11.7*	*2.3*	*0.0*
Si3–Si4	2.374	0.0	8.54	−12.2	2.3	0.0
	2.375	*0.0*	*8.53*	*−12.1*	*2.3*	*0.0*
Bulk Si						
Si–Si	2.375	0.0	8.53	−12.1	2.3	0.0
SiH_4						
Si–H	1.476	3.8	11.9	37.3	80.1	0.0

[a] Ref. [42]; RHF/3-21G(*d,p*) periodic or molecular wavefunctions; slabs: totally relaxed geometrical/basis set model [42]. If not otherwise stated, all quantities in a.u.

[b] Data for the clean surface in *italics*; Δ_{BCP} is the distance from the BCP to the X–Y bond midpoint, expressed as a percentage of half the bond length. A positive Δ_{BCP} value indicates that the BCP is closer to X.

tion the "surface layer", i.e. that portion of a crystal having nonbulk properties, is referred to as surface *double layer*. The "surface layer" of the hydrogen-covered system may be described as a double layer in which the internal layer is composed of Si1 and Si2 atoms. The H and Si1 atoms in the hydrogen-covered surface polarize in such a way as to oppose the electric field created by the CT within the surface double layer. In fact, the H and Si1 atomic dipole nonzero components μ_z are both directed away from the surface (Table 7.6). The hydrogen coverage has the effect of reversing and greatly enhancing the magnitude of the Si1 polarization in the clean system, in which Si1 accumulates negative charge outwards. This agrees with the differences between the Laplacian distribution for the two slabs (Fig. 7.4 bottom). Four-bonded charge concentrations (BCCs) along the corresponding Si–Si bonds are observed for all the Si atoms, except Si1, in both systems. Si1 in the hydrogen-covered surface lacks the BCC associated with the Si1–H bond, because of the large amount of electron transfer from Si1 to the H atom and the ensuing loss of an Si VSCC along this bond. Instead, in the clean surface, the dangling sp^3 orbital per surface atom manifests itself as a nonbonded charge concentration (NBCC), characterized by a closer distance from the Si nucleus and half of the Laplacian value at a BCC in the bulk.

By analogy with the bonding properties, almost perfect convergence toward the bulk values are also observed for the integrated properties beyond the first two or

Table 7.6 Relationship between atomic properties and depth from surface of hydrogen-covered and clean Si(111)(1×1) slabs compared with silicon bulk.[a,b]

Ω	q(Ω)	μ_z(Ω)	ΔE(Ω)	ΔV_1(Ω)	*D*(Ω)
Si(111)(1×1):H; *Si(111)(1×1) clean*					
H	−0.736	+0.43			
Si1	+0.698	+1.23	0.176	−18.1	3.4
	+0.003	*−0.31*	*−0.206*	*+42.8*	*8.5*
Si2	+0.053	−0.03	0.011	+0.1	2.3
	−0.019	*+0.06*	*0.068*	*+3.5*	*1.9*
Si3	−0.015	−0.03	0.016	−0.4	0.8
	+0.013	*+0.02*	*0.066*	*−1.4*	*0.5*
Si4	−0.002	−0.01	0.017	−1.0	0.7
	+0.001	*+0.01*	*0.060*	*−0.5*	*0.6*
Bulk Si					
Si	0	0	0	0	0.6

[a] Ref. [42]; wavefunctions and slabs model as for Table 7.5. All quantities in a.u.; data for the clean surface in *italics*.
[b] The only nonzero μ_z component of the atomic dipole normal to the surface is reported. The z axis is directed away from the surface. Energies ΔE and volumes ΔV_1 are given relative to bulk silicon ($\Delta X = X_{slab} - X_{bulk}, X = E, V_1$); diffuseness, D, is given by $D = [(V_1 - V_2)/V_1] \times 100$.

three atomic layers. The only exception is the atomic energies, which is not found to converge to bulk values in both slabs, because the electric field created by the surface double layer shifts the binding energies of electrons. This is, therefore, the result of a physical phenomenon known to occur in any finite crystal, i.e. a crystal having surfaces [45]. A more detailed explanation is given on page 895 of Ref. [42].

Figure 7.4 (middle) shows that in the clean slab both Si1 and Si2 atoms have infinite volume, whereas starting from the Si3 atoms inward the atoms have a finite volume, as in the bulk. The surface nature, in particular that of the Si1 atoms, is made evident by their increased V_1 volume, relative to that in the bulk (<30% for Si1). In contrast, in the hydrogen-covered slab the Si1 atoms have a finite volume, largely shrunk relative to that in the bulk, and are completely isolated from the outside because of their interaction with H atoms. In this slab the Si2 atoms become the only, nonfinite volume, surface Si atoms, an observation which explains why, experimentally, passivant substitution or oxidation are mediated by Si2 atoms and never occur directly at Si1 atoms [46].

An interesting measure of the "surface" character of an atom in a slab may be obtained by defining a diffuseness D of its electron distribution, given by $D = [(V_1 - V_2)/V_1] \times 100$. As expected, the largest value of D is found for the

Si1 atom in the clean slab (Table 7.6). This atom is characterized by an electron density lower than 0.002 a.u. in more than 8% of the volume of its V_1 basin, whereas the corresponding values for Si bulk or for atoms Si3 to Si5 in both slabs do not exceed 0.8%.

7.4.2 Si(111)(2×1) Reconstructed Surface

The Si(111) surface structure obtained after low-temperature ($T < 600$ K) ultra-high vacuum cleavage, has a (2×1) reconstruction [39], with two buckled layers of chain-bonded surface atoms (Fig. 7.4). This surface is a prototype of clean semiconductor surfaces and, although apparently simple, has several intriguing features which have motivated almost 30 years of experimental and theoretical efforts [39, 41, 44]. The top layer is formed by threefold-coordinated atoms (Si1 and Si2) and the lower layer by fourfold-coordinated atoms (Si3 and Si4). (There are two unique atoms per layer in the reconstructed surface and one unique atom per layer in the non reconstructed surface. The unique atoms of the first layer (Si1, Si2) and second buckled layer (Si3, Si4) are, respectively, in ideal correspondence with the Si1 and Si2 unique atoms in the unreconstructed Si(111)(1×1) surface. A centrosymmetric slab, with a 2×1 unit cell and consisting of 14 layers, for a total of 28 Si atoms per cell was used [43] to model the (111)(2×1) surface (Fig. 7.4, top, right).) On reconstruction the surface dangling bonds become located on nearest-neighbor sites rather than on next-nearest neighbor sites as in the (111)(1×1) surface and the reconstructed surface structure is customarily interpreted in terms of the Pandey's π-bonded chain model [47]. According to this model, the zigzag chain of adjacent p_z orbitals can, in principle, π-bond as in organic materials, forming bonding and antibonding surface π states.

QTAIM is here used to quantify how close to reality is Pandey's interpretation of the driving force under surface reconstruction [43]. Analysis of the Laplacian distribution (Fig. 7.4, bottom), shows that in the first buckled layer only Si1 and not Si2 atoms have an NBCC pointing away from the surface. The Si1 NBCC, moreover, has a magnitude which is approximately 20% less than that found for Si1 in Si(111)(1×1). After reconstruction, therefore, the single electrons associated with the nominal dangling bonds of the threefold-coordinated atoms either (Si2) take part entirely in bond reconstruction or (Si1) become at least partially involved in rebonding, compared with the first layer atoms in the Si(111)(1×1) surface. Inspection of bond ellipticity profiles along the bond paths of each pair of bonded atoms in the two outermost layers of the reconstructed surface adds further insight (Fig. 7.5). As expected, the highest π character is that of the Si1–Si2 bond in the first layer. The ellipticity profile along the bond path is indicative of a very asymmetric contribution to the π bond from the two atoms, in accordance with the large differences found in their Laplacian distributions and atomic properties [43]. Interestingly, significant ellipticities are also observed for the bonds between the first and second layer (Si2–Si3, Si1–Si4) and those within the second layer (Si3–Si4). This reveals that the π conjugation extends over a 2D

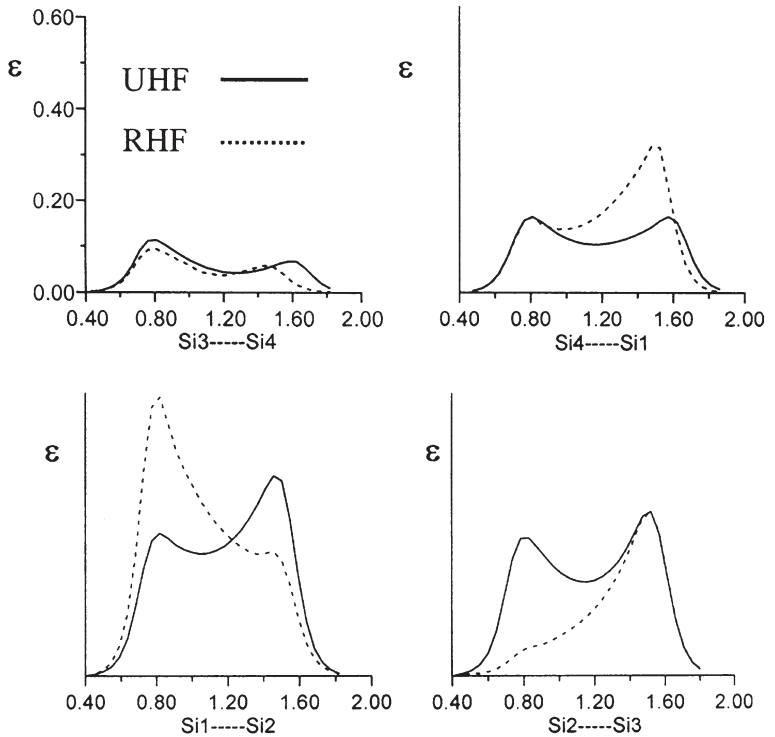


Fig. 7.5 Si(111)(2×1) surface. Bond ellipticity (ε) profiles along the bond path of each unique pair of bonded atoms in the two outermost layers. Results refer to the LEED geometry for the surface [48]. Adapted, with permission, from Fig. 4 in Ref. [43], F. Cargnoni, C. Gatti, *Theor. Chem. Acc.* 2001, *105*, 309–322 (Copyright 2001, Springer).

array of bonds, rather than being strictly localized along the topmost layer 1D chains, as hypothesized in the Pandey's model. Adoption of a spin-polarized RHF solution, in place of a simple RHF picture, leads to a significant energy reduction for the reconstructed surface, accompanied by an increased atomic equalization within each pair of bonded atoms and by noticeable enhancement of the surface bonds π-delocalization (Fig. 7.5).

In general, the (2×1) reconstruction involves much larger surface-cell charge distortions than found in the (1×1) surface and noticeable changes in the atomic polarization of the surface layer atoms. The effect of buckling is to largely differentiate the properties of the two unique atoms of each surface layer. For example, the net charge of Si1 atom is $+0.138e$ and that of Si2 is $-0.225e$ (UHF results, at LEED geometry), compared with the value of $+0.003$ e for Si1 in the (111)(1×1) surface. These charge rearrangements, along with the onset of a 2D partially π-delocalized bond network, lead to Si–Si bonds in the first two layers that are, on average, as strong as in the crystal [43]. This contrasts with the non reconstructed

surface in which the corresponding Si1–Si2 bonds are, instead, weaker than in the bulk (Table 7.5).

Comparison of the Si1 and Si4 basins (middle, right) with the Si1 and Si2 basins (middle, center) in Fig. 7.4 shows that after the 2×1 reconstruction only the atoms of the topmost layer (Si1, Si2) have infinite volume whereas the atoms of the second layer (Si3, Si4) – differently from those in the unreconstructed surface – are already finite, as in the bulk. Convergence toward bulk values of both bond and atomic properties was found to start from the fifth layer inwards, i.e. at a greater depth from surface than in the clean and hydrogen-covered (111)(1×1) surfaces. The reader is referred to Ref. [43] for full details of this QTAIM study.

7.5
QTAIM Applied to Host–Guest Systems

Host–guest chemistry is a fundamental branch of supramolecular chemistry [49] and the guest@host concept has served to describe the structure of most supramolecular systems [50]. Typical questions of host–guest chemistry are the nature of bonding interactions between the guest and the host and the amount of electronic charge transferred to the host from the guest, or vice versa. These are not just academic issues, because the exciting functional properties of these systems, and their optimum tuning, may crucially depend on how the guest binds to the host and how this binding may be suitably chemically modified. Standard methods of bond analysis and charge partitioning may prove particularly inadequate in this area, because these supramolecular assemblies are usually held together by noncovalent interactions (HBs, van der Waals forces, π–π interactions, and/or electrostatic effects), which are much weaker than covalent bonds and, except for HBs, usually nondirectional. QTAIM, being firmly rooted in physics and not based on any preconception of bonding, is once more the theory of choice.

QTAIM is applied here to two peculiar host–guest crystalline systems – the type I inorganic clathrates $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba), which have promising thermoelectric properties [51], and sodium electrosodalite (SES), in which the guest, which explicitly requires QTAIM to be identified and localized, gives rise to an ordered bcc lattice of F-centers with concentration three order of magnitude higher than in the ionic solids [52].

7.5.1
Type I Inorganic Clathrates $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba)

Type I clathrates consist of two type of cage, the twenty-atom dodecahedron and the twenty-four-atom tetrakaidecahedral cages (Fig. 7.6, left), which are usually formed by group 13 and Group 14 elements and which each encapsulate a guest metal atom A [53]. The metal atoms ''rattle'' in these oversized cages and are known to have localized, low-frequency phonon modes that reduce the thermal conductivity to values comparable to those of amorphous semiconductors, with-

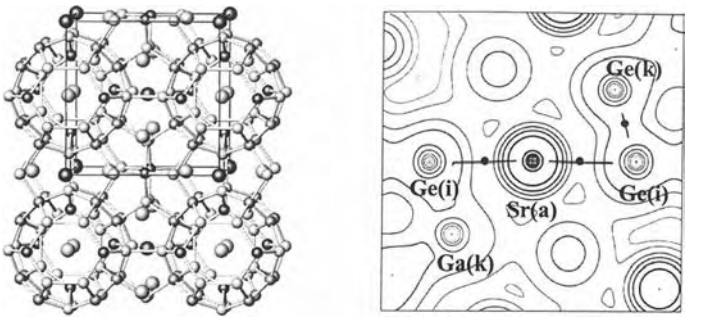


Fig. 7.6 $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba) (a) The clathrate type I structure (fully symmetric ideal structure, $Pm\bar{3}n$ space group). The large dark atoms are A guests (2*a* site) hosted in the twenty-atom cages and the large gray atoms are A guests (6*d* site) hosted in the twenty-four-atom cages. The small dark gray and light gray atoms are the 6*c*, 16*i* and 24*k* framework sites, respectively. In the $Pm\bar{3}n$ structure Ga atoms are located at 16*i* positions. (b) $Sr_8Ga_{16}Ge_{30}$ at minimum P1 geometry – contour map of the ab-initio periodic pseudo-potential (PP) density in a plane containing an Sr atom in a twenty-atom cage (*a* site) and two Ge atoms at positions *i* and *k* at distances of 3.46 and 3.65 Å from the Sr atom. The nuclei on the left of the central Sr atom are less than 0.1 Å from the plane of the plot. Positions are labeled according to $Pm\bar{3}n$ space-group. The bond paths do not terminate at nuclei, because this is a PP density. The Ge and Ga atoms at *k* positions, with distances from Sr greater than 3.6 Å turn out to be nonbonded to the central Sr atom in the P1 geometry relaxed structure. Adapted, with permission, from Figs 1 and 4 of Ref. [59], C. Gatti, L. Bertini, N. P. Blake, B. B. Iversen, *Chem. Eur. J.* 2003, *9*, 4556–4568 (Copyright 2003, Wiley).

out significantly affecting the semiconductor-like properties of the host cages' crystalline framework [51]. These systems thus have, potentially, the ideal thermoelectric properties of a phonon glass and an electron single crystal (PGEC) [54].

Structural chemists usually regard these clathrates as Zintl phases in which the guest atoms completely transfer/accept valence electrons to/from the framework [53]. This belief is firmly supported by the observation that all known clathrate type I structures have a common number of 184 valence electrons per unit cell, despite the large number of elemental compositions forming these structures.

Studies on the very promising $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba) thermoelectrics based on theoretical EDD [55] or maximum entropy method (MEM) analysis of experimental EDD [56] have, however, called into question the ionic character of the guest atoms and rather propped up the idea of their almost neutrality. The guest atoms are usually found far from the cage center and with displacements and vibrational frequencies which differ from guest to guest and, for a given guest, from the large to the small cage [55, 56, 57]. This experimental observation gives some credibility to the idea of a not fully ionic metal guest, because a directional, not purely electrostatic guest–host interaction, with possible incomplete CT, seems to be operating. Analysis of the density of states (DOS) and bands features obtained from the same wavefunction from which the theoretical EDD was

Table 7.7 $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba): Atomic net charges of guest atoms A compared with corresponding charges in SrO and BaO.[a,b]

System/model	Ω	q(Ω)/e	
		Small cages	Large cages
Clathrate/ab initio	Sr	1.669 (0.001)	1.687 (0.005)
	Ba	1.743 (0.001)	1.773 (0.003)
Clathrate/IAM	Sr	−0.096	−0.126
	Ba	0.858	0.576
SrO/ab initio	Sr	1.717	
BaO/ab initio	Ba	1.758	

[a] Ref. [59], data for the most stable P1 clathrate structures [55] and for the SrO and BaO crystals, T2 wavefunction model [59]. There are 2 and 6 unique twenty-atom and twenty-four-atom cages in the P1 structure; average net charges (standard deviations) are listed.
[b] IAM computation at the $Pm\bar{3}n$ space group geometry.

derived strongly suggests, however, that the guest atoms are acting as electron donors [55, 58], which seems in sharp contrast [59, 60] with their claimed almost neutrality.

We used QTAIM to shed light on these conflicting views. Table 7.7 shows QTAIM net charges for the guest atoms A for the most stable P1 space group-optimized structure of each clathrate [58]. Irrespective of whether they occupy the large or the small cage, guest atoms are almost completely ionized. Their net charges, averaged over the two and the six unique twenty-atom and twenty-four-atom cages in the crystal cell, are approximately 1.68*e* and 1.77*e* for the Sr and Ba clathrates, respectively. These values are to be compared with the corresponding charges of 1.72*e* and 1.76*e* in the SrO and BaO oxides, which are both well known examples of crystalline ionic compounds. Very similar net charges for A = Sr, Ba are obtained for the second more stable clathrate structure and also for the idealized, fully symmetric $Pm\bar{3}n$ structure having all the Ga atoms in the 16*i* framework site (Fig. 7.6, left).

The question arises of why this clear-cut picture of the guest ion/neutral atom dilemma is so at odds with previous suggestions from theory [55] and experiment [56]. The main reason is that in these studies charge transfer (CT) was defined relative to a *reference density*, rather than using QTAIM to identify the actual shape of the atomic basins in the clathrate and then obtain their electron populations. The reference density was taken as the density of the interleaved, noninteracting A_8 and $Ga_{16}Ge_{30}$ sublattices in the theoretical study [55], or as the nonuniform prior density in the MEM experimental analysis. In both studies the electron population on a guest atom, calculated by integrating the EDD within a sphere of given radius centered on the atom, was found to be basically constant whether

computed from the true density of the clathrate or from the reference densities detailed above. Hence guest-to-host charge transfer was inferred not to occur in the clathrate. Such an empirical evaluation differs substantially from the QTAIM result [59], because when the interaction among the host framework and the guest atoms is turned on, new attractive and repulsive physical forces come to the play and the EDD should no longer be geometrically apportioned among the host and the guest as when the interaction is turned off. Table 7.7 shows that QTAIM partitioning of the IAM density, as opposed to that of the crystal density, would predict much less ionized guest atoms – the Sr atoms becoming almost neutral and the Ba atoms having their net charge more than halved. The IAM density is not a quantum object, nor is there any physical need for the zero-flux recipe to predict the A atoms to be neutral for IAM densities. It is, however, worth noting that the claim of the presence of almost neutral guest atoms in $A_8Ga_{16}Ge_{30}$ roughly corresponds to the outcome one would obtain if the QTAIM analysis, instead of being applied to the true density of the clathrate, were to be performed on a model density which neglects the charge rearrangement because of bonding.

By describing the guest atoms as highly ionized species, QTAIM analysis reconciles theory with the Zintl phase view structural chemists have of the inorganic class I clathrates. QTAIM results also agree with the negligible contributions provided by Sr and Ba to the valence states close to the Fermi level and with the dominant contributions Sr and Ba provide to the conduction states close to this level (see DOS analysis in Ref. [59]).

Geometry relaxation from the fully symmetric $Pm\bar{3}n$ clathrate structure to the most stable and second most stable P1 structures leads to noticeable displacement of Sr and Ba atoms from the centers of the twenty-atom and, particularly, the twenty-four-atom cages, for which the displacement is as large as 0.81 and 0.42 Å for Sr and Ba, respectively [56]. The force driving this guest atom disorder lies in the replacement of the many, very weak, guest–host bonds present in the fully symmetrical clathrate cages (twenty for the small cage and eight for the large cage) with the fewer, stronger and shorter chemical interactions characterizing the P1 geometry clathrates (Fig. 7.6 right) [59]. For example, only 2.8(8) and 4.0(11) guest–host BCPs are found on average for the Sr and Ba clathrates in the large cages, instead of the eight present in the fully symmetrical structure. These remaining bonds are, however, substantially stronger, as indicated by the 50% (Sr) and 16% (Ba) average increases of their ρ_b values relative to those of the corresponding bonds in the $Pm\bar{3}n$ clathrate structure. Differences between the rattling frequencies in the Sr and Ba clathrates, and between their larger or smaller cages, have recently been related to, and found to comply with, the relative strength of the guest–host interactions, as ordered by the ρ_b values [57]. On the whole, all the recovered bonds, including those significantly strengthened after clathrate geometry relaxation, retain properties typical of a closed-shell ionic interaction (low ρ_b, low and positive $\nabla^2\rho_b$; BCP location close to the positive ion), with the guest–host binding energy being essentially electrostatic in nature. Full details on this study may be found in the original papers [59, 60].

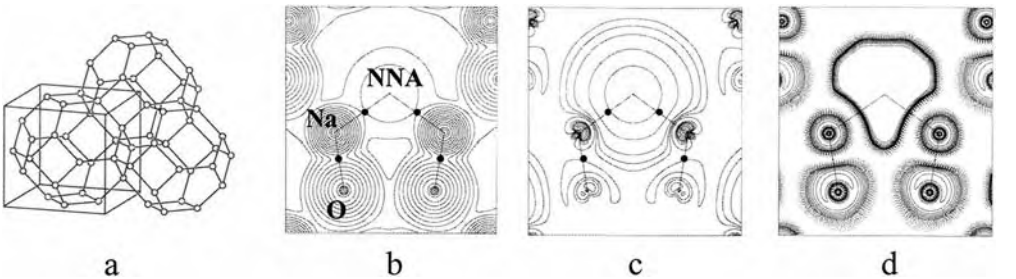


Fig. 7.7 *F* center in sodium electrosodalite (SES). (a) Schematic diagram of the sodalite framework formed by a bcc array of β cages consisting of regularly alternating SiO_4 and AlO_4 tetrahedrons. The circles in the figure correspond to regularly alternating Al and Si atoms bridged by an O atom. In the SES (space group $P\bar{4}3n$), the Na_4^{3+} tetrahedral clusters are located inside each β cage. There are two β cages and two unpaired electrons in each unit cell (ferromagnetic phase). The NNAs associated with the *F* centers are located at the center (2*a* site) of the Na_4^{3+} clusters and form a bcc lattice, with same cell parameter as SES. (b) Total electron density and (c) spin density in the sodalite cage, in the $(1,1,0)$ plane. Contour levels at 2, 4, and 8×10^n a.u., with *n* ranging from 0 to -3 for the total density and from 0 to -4 for the spin density. In the spin density the first line is the zero contour. The contour level closest to the NNA has similar shapes and equal value (0.004 a.u.) in both maps, because the *F* center contains almost solely unpaired electron density. (d) ELF^{α} in the $(1,1,0)$ plane. The first dashed contour is the zero contour and the increase is 0.05 per contour up to 0.45. The first full contour has the value 0.5 and the increase is 0.05 per contour up to 1.00. (Adapted, with permission, from Figs. 1, 2, and 4 from Ref. [62] G. H. K. Madsen, C. Gatti, B. B. Iversen, L. Damjanovic, G. D. Stucky, V. I. Srdanov, *Phys. Rev.* 1999, *B59*, 12359–12369; Copyright 1999, American Physical Society).

7.5.2
Sodium Electrosodalite

The sodalite host framework, $(SiAlO_4)_6$, is a bcc array of β cages consisting of regularly alternating SiO_4 and AlO_4 tetrahedrons (Fig. 7.7a). To balance the formal negative charge of the host, each β cage contains three positive guest Na atoms. When exposed to sodium vapor the sodalite gradually turns blue, then purple, and eventually black [61]. The color change has been ascribed to formation of *F* centers in the tetrahedral Na_4^{3+} clusters containing the three existing Na ions and the excess Na atom absorbed in each β cage. The resulting sodalite structure, named sodium electrosodalite (SES) has cubic symmetry with each unit cell containing two β cages and two perfect tetrahedral arrangements of sodium atoms inside each β cage $[Na_8(SiAlO_4)_6]$. The unit cell thus contains two unpaired electrons and SES undergoes an antiferromagnetic transition at 48 K [52]. The unpaired electrons are thought to have predominantly *s* character, because of the spherical symmetry of the sodalite cage, and this material has been identified as the first example of an *s*-electron antiferromagnet. Electron density distributions for the ferromagnetic and the antiferromagnetic phase are very much alike, with

Table 7.8 Atomic and nonnuclear attractor (NNA) properties in the ferromagnetic phase of sodium electrosodalite $Na_8(AlSiO_4)_6$.[a]

Ω	N(Ω)	$N_{\alpha-\beta}(\Omega)$	$V_T(\Omega)$	$G(\Omega)/N(\Omega)$
NNA	0.730	0.692	316.3	0.09
O	9.757	0.012	135.5	7.69
Na	10.095	0.041	71.9	15.93
Si	10.572	0.000	18.2	27.31
Al	10.369	0.000	23.3	23.13

[a] Ref. [62]; all quantities in a.u.; $N_{\alpha-\beta}$ is $(N_\alpha - N_\beta)$, the excess number of electron of spin α in the atomic basin and $G(\Omega)/N(\Omega)$ is the basin kinetic energy G per electron.

alternating ordering of the spin density being the only difference between the two phases [62]. (For this reason, we report here only the QTAIM analysis for the ferromagnetic phase.)

When the guest Na atom is added to the sodalite host, we know exactly, from structural analysis, where the Na cation goes. It becomes totally indistinguishable from the pre-existing Na cations. The question remains, however, of where the extra electron added to each sodalite cage ''locates''. Is it primarily delocalized over each Na ion clusters or is it (well) localized somewhere? And, supposing we can physically ''identify'' this electron as another ''atomic'' constituent of the lattice, can this *new* constituent be related to the formation of the bcc lattice of *F* centers in SES?

As shown in Fig. 7.7b, QTAIM analysis of the SES periodic UHF electron density distribution finds an NNA, located at the center (2*a* site) of each Na_4^{3+} cluster and bonded to its neighboring four Na atoms [62]. Spatial distribution of NNAs gives rise to a bcc lattice with the same cell parameter as SES and with the same spatial distribution of the *F* centers determined from spectroscopy. The unpaired electron therefore has a separate identity – it behaves as a quantum mechanical open system and has its own set of properties (Table 7.8). In the real space, the *F* center manifests itself as the union of a maximum in the electron density at the nonnuclear position 2*a* and of the basin associated with this NNA. As shown in Table 7.8, the *F* center has an electron population of 0.730*e* and contains almost solely unpaired electron density, with over 97% of the electrons being spin α electrons. This explains why the electron density and the spin density distribution in the sodalite cage look very much the same in a large region surrounding the NNA (Figs. 7.7b and 7.7c). Over 69% of the unpaired density in the unit cell is contained in the two *F* center basins, the remaining 31% being shared between the eight Na (16.4%) and the twenty-four O basins (14.4%).

We now use the ELF distribution for the α-spin orbitals, ELF^α [21, 62], to demonstrate that the unpaired electron density in the *F*-center basin is to a large extent associated with a *localized* α-spin *orbital*. Indeed, the Na–NNA BCP lies close

to the 0.5 ELF^{α} contour (Fig. 7.7d) and, by moving toward the center of the Na tetrahedron, i.e. toward the NNA, ELF^{α} rapidly increases regularly up to the value of unity. If the density is dominated by a single, localized α-spin orbital in the NNA basin, the local curvature of the spherically averaged same-spin conditional pair density has to vanish or nearly vanish [10] and, as a consequence, the ELF value must approach unity. As expected, complementary behavior is observed for the corresponding map for ELF^{β} [21, 62], with most of the volume of the NNA basin being characterized by an ELF^{β} value close to zero.

The NNA basin has the largest basin volume (316 a.u.) and a very low and flat electron density, yielding an average volume per electron of 433 a.u., approximately 31 times than that of the oxygen anions, the second most "diluted" atoms. The 1.46*e* of the two *F*-center basins in the unit cell occupy 13.4% of the crystal volume, even though they represent only 0.32% of the total electronic charge. The electron density is very loosely bound in the *F*-center basins, as indicated by the very low kinetic energy per electron, $G(\Omega)/N(\Omega)$ value of 0.09 a.u. (Table 7.8). Such a value can be compared with those of other NNA basins associated with "free electrons" in the alkali metal clusters (0.06–0.08 a.u.) [17, 62, 63], in the Be metal (0.282 a.u.) [62, 64], and in a cluster model of an *F*-center in an ionic solid (0.184 a.u.) [65].

7.6 The Source Function: Theory

Few years ago, Richard Bader and I showed [14] how the electron density at any point $\mathbf{r}$ within a molecule may be viewed as consisting of contributions from a local source $LS(\mathbf{r}, \mathbf{r}')$ operating at all other points of the space where the local source has the expression:

$$\rho(\mathbf{r}) = \int LS(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' = \int_{\Omega} LS(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' + \sum_{\Omega' \neq \Omega} \int_{\Omega} LS(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' \tag{3}$$

$$LS(\mathbf{r}, \mathbf{r}') = -(1/4\pi)\frac{\nabla^2 \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{4}$$

In Eq. (4), the Green's function $(4\pi|\mathbf{r} - \mathbf{r}'|)^{-1}$ is an *influence function* [66], which represents the effectiveness of the *cause* $\nabla^2\rho(\mathbf{r}')$ to give rise to the *effect* $\rho(\mathbf{r})$. The effectiveness depends on the reciprocal of the distance between the element of Laplacian of the density $\nabla^2\rho(\mathbf{r}')\, d\mathbf{r}'$ and the point of interest given by $\mathbf{r}$. (By combining Eqs. (3) and (4), one gets $\rho(\mathbf{r}) = -(1/4\pi)\int \frac{\nabla^2\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$. This expression is formally equivalent to that for $W(\mathbf{r})$, the electrostatic potential generated at $\mathbf{r}$ by the system's electron distribution $\rho(\mathbf{r})$. One may envisage $\rho(\mathbf{r})$ as the potential generated at $\mathbf{r}$ by the system's Laplacian density distribution [14].)

Calling the integral of $LS(\mathbf{r}, \mathbf{r}')$ over the basin of an atom or group of atoms the source function (SF) contribution, $S(\mathbf{r}, \Omega)$, from that atom or group of atoms to $\rho(\mathbf{r})$:

$$\int_{\Omega} LS(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' \equiv S(\mathbf{r}, \Omega) \tag{5}$$

one can envisage the electron density at a point within an atom as determined solely by an internal SF self-contribution and by SF contributions from the remaining atoms or groups of atoms within a molecule:

$$\rho(\mathbf{r}) = S(\mathbf{r}, \Omega) + \sum_{\Omega' \neq \Omega} S(\mathbf{r}, \Omega') \tag{6}$$

The SF is thus a measure of the relative importance of an atom's or group's contribution to the density at any point [14]. The decomposition afforded by Eq. (6) enables one to view the properties of the density from a new perspective and establishes the SF as a novel tool for providing chemical insight [14, 67]. (Modified versions of the PROMEGA/PROAIMV and TOPOND-98 software enable one to evaluate the atomic SF contributions to the density at a selected list of points. This software is available from the author upon request.)

At this point one may ask oneself which is the physical meaning behind the local source $LS(\mathbf{r}, \mathbf{r}')$. Several interpretations are possible [14, 66, 68], the more insightful being probably that which links $LS(\mathbf{r}, \mathbf{r}')$ with the local expression of the virial theorem [6]:

$$LS(\mathbf{r}, \mathbf{r}') = -\frac{1}{\pi} \frac{2G(\mathbf{r}) + V(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \tag{7}$$

The local source is thus related to the failure to locally satisfy the virial relationship between twice the integrated kinetic energy and virial field densities. Molecular regions where the electron density is concentrated ($\nabla^2\rho(\mathbf{r}') < 0$) and where the potential energy dominates the kinetic energy are a *source* for the electron density at a point $\mathbf{r}$. Conversely, regions where the electron density is depleted ($\nabla^2\rho(\mathbf{r}') > 0$) and where the kinetic energy dominates potential energy act as a *sink*, removing electron density at $\mathbf{r}$. The effectiveness of the electron density at $\mathbf{r}'$ to be a source or a sink for the electron density at another point $\mathbf{r}$ is then related to the magnitude of its charge concentration or depletion at $\mathbf{r}'$, weighted by the inverse of the distance between the two points.

The SF may be investigated using as a reference point any point $\mathbf{r}$ in a system, including the nuclei and the NNAs [67]. The BCPs have usually been taken as the least biased choice for points representative of bonding interactions [14, 67]. For an isolated atom, the SF contribution to any point $\mathbf{r}$ is always positive, because $\rho(\mathbf{r})$ is ≥ 0 everywhere. For systems with more than one atom, it is found that, in

general, the positive LS contributions from an atomic basin to the density at system's BCPs dominate its negative contributions. In some special circumstances, however, it may happen that an atom acts as a sink rather than as a source for the density at a given system's BCPs [67]. The relevance of such instances will be discussed in Sections 7.6.2 and 7.6.3.

Analysis of $LS(BCP, \mathbf{r}')$ profiles with $\mathbf{r}'$ running along a bond path have been discussed for the series of second-row diatomic hydrides and for several hydrogen-bonded systems in the gas and crystalline phases [68]. With respect to the integral form of the SF function, the study of its local form, $LS(BCP, \mathbf{r}')$, along a bond path, introduces further detail. It enables discovery of which regions in the basins of the two linked atoms accumulate or remove charge at the BCP, as a function of the nature of the bond and of the external perturbations (substituent effects, crystal field, etc.) acting on this bond. The local form of the SF is discussed no further in this chapter. It does, however, represent an interesting fingerprint of a given bonding interaction and the reader is referred to the original paper [68] for further details.

7.6.1
The Source Function and Chemical Transferability

The SF has been used in several examples [14, 67] as a very sensitive measure of an atom's or chemical group's transferability and of the consequences thereof. Indeed, the "perfect" transferability of a group property from one molecule to another not only implies a corresponding transferability of the group's electron density but also that the sum of contributions to this density from the remaining atoms or group of atoms in the system remain constant.

After ethane the terminal methyl group in *n*-alkanes is known to have transferable atomic properties, for example energy, electron population, volume, and spectroscopic properties, irrespective of the length of the chain [6]. The transferability of the electron distribution in the methyl group is good enough also to yield a constant value of ρ_b at its unique C–H bond. The SF group contributions to this ρ_b value are displayed in Table 7.9 for ethane, propane, butane, and pentane [14]. The contribution from the atoms in the methyl group to ρ_b are constant at 0.270 a.u. throughout this series, with the contribution from the two equivalent hydrogen atoms of the methyl group to this amount equaling 0.0210 a.u. for all four molecules. The contribution from the neighboring methylene group seems to be constant after ethane. The constancy in the ρ_b value (0.283 a.u.) for the terminal C–H bond is because there is a constancy in the sum of the source contributions to ρ_b from groups external to the methyl group, the quantity $S(\mathbf{r}_b; \text{ext})$, *irrespective of the length of the chain*. Thus extending the propane chain by removing a hydrogen atom from the external methyl group and replacing it with another methyl group yields an ethyl group, whose SF in butane contributes 0.0036 a.u. to ρ_b compared with 0.0035 a.u. from the external methyl group in propane. When this ethyl group is further extended by a methylene group to ob-

Table 7.9 Source function contributions to H–CH_2 BCP in methyl group.[a]

Molecule	Atomic group source function contributions					ρ_b(H–CH_2)	S(r_b; ext)
Ethane	H–CH_2	CH_2	H			0.2830	0.0126
	0.2704	0.0100	0.0026				
Propane	H–CH_2	CH_2	CH_3			0.2827	0.0126
	0.2701	0.0091	0.0035				
Butane	H–CH_2	CH_2	CH_2	CH_3		0.2827	0.0127
	0.2701	0.0091	0.0020	0.0016			
Pentane	H–CH_2	CH_2	CH_2	CH_2	CH_3	0.2827	0.0127
	0.2702	0.0090	0.0019	0.0008	0.0009		

[a] Ref. [14]; all quantities in a.u.; S(r_b; ext) is the sum of source contributions to ρ_b from groups external to the methyl group. It is given by the sum (except H–CH_2) of the atomic group source contributions listed in the second column for each molecule.

tain the propyl group in pentane, the SF contribution remains unchanged at 0.0036 a.u. Thus the ethyl group in propane contributes the same as the butyl group in pentane, with the result that the H and C atoms of the methyl have their characteristic properties in hydrocarbons.

The series Li–X (X = F, O, N, Cl, H) is another interesting example of the use of the SF as a very sensitive tool to test the extent of chemical transferability [67]. The Li atom in this series is known to have nearly constant and transferable properties, including its net charge and atomic energy [6]. In the limit of perfect transferability, the SF contribution from Li at the BCP should remain constant along the series and changes in ρ_b values, as a function of X, should only be determined by correspondingly equal changes in $S(\mathbf{r}_b, X)$. As shown in Ref. [67], the ρ_b value decreases by almost a factor of two along the series and similar behavior is found for $S(\mathbf{r}_b, Li)$. The change of $S(\mathbf{r}_b, Li)$ along the series indicates that this quantity is a more sensitive index of a departure from perfect transferability than are the integral averages yielding the atomic population and energy. Interestingly, the profiles of $S(\mathbf{r}, Li)$ along the Li–X axis, show that the Li atom makes a constant contribution to $\rho(\mathbf{r})$ for any X, up to a distance of 0.58 Å from the Li nucleus (Fig. 3 in Ref. [67]). This is near the distance of the BCP closest to the Li nucleus in the series. The observed changes in $S(\mathbf{r}_b, Li)$ are merely a consequence of the progressive shift towards the X nucleus along the series.

The most important result is that the contributions from Li to ρ_b remain almost constant throughout the series, at approximately 40%, however. The constant cationic nature of Li in the LiX series seems to be mirrored in a constant percentage SF contribution from Li to ρ_b rather than in a constant source. The shift in the BCP location along the series serves to maintain constant this relative contribution of Li to ρ_b.

7.6.2
Chemical Information from the Source Function: Long and Short-range Bonding Effects in Molecular Complexes

This section shows how the SF may be used as a tool to disclose the extent of local/nonlocal character of a bonding interaction, using only information contained in the system's electron distribution. Hydrogen-bonded molecular complexes [67] are discussed at length and very recent results [69] on the "metal–metal" bond in polynuclear metallic complexes are discussed briefly.

Energies of hydrogen-bonds, which extend from approximately 15 to 50 kcal mol^{-1} for strong bonds down to 1 to 4 kcal mol^{-1} for weak bonds, are evidence of a wider range of atomic interactions than is observed for covalent or ionic bonds or van der Waals forces [70]. The dominating energy contributions depend on the nature of the H-donor (D) and H-acceptor (A) atoms and of the distance, $d_{D\cdots A}$ between them.

The reaction path for approach of two water molecules, within the linear C_s constraint, is first introduced as a guide to understanding how the SF contributions to the hydrogen bond ρ_b value change with changes in the nature of the HB. Table 7.10 reports percentage atomic sources, $S(\Omega' + \Omega'' + \cdots)\%$, from $d_{D\cdots A}$ distances typical of a weak isolated HB to those occurring in very strong HBs, the $d_{D\cdots A}$ equilibrium distance being 3.020 Å, at the RHF/6-311G(2d,2p) level. As shown in the table and in Fig. 7.8a, the atomic percentage contributions change dramatically along the reaction path. It is only the percentage global contribution from either the H-donor or the H-acceptor molecule which remains almost constant, and not dissimilar from each other, at values of approximately 44–48% and 56–52%, respectively. Also stable enough, and approximately 30–40%, is the value of S(H+D)%, the percentage contribution from the H involved in the HB and the oxygen donor D. Its relative constancy indicates that, besides the donor and acceptor molecules, D–H also behaves to a some extent as an atomic "transferable" group. Conversely, the percentage source contribution from the H, S(H)% seems to be the most distinctive marker of the change of the nature of the hydrogen-bond along the reaction path. S(H)% is very small and positive only for very short $d_{D\cdots A}$ distances and rapidly becomes negative and even highly negative at distances larger than 2.5 Å. At equilibrium distance, S(H)% is as negative as −72.3%. Why is this so? In weak and moderate strength HBs, the shape of the H basin and of the Laplacian distribution along the OH···O axis are highly asymmetric, with regions of negative Laplacian surrounding the O–H BCP and with the HB critical point being, instead, located in a region of pronounced positive Laplacian (Fig. 6, Ref. [67]) and, hence, of negative local source contributions, $LS(BCP, \mathbf{r}')$. As the $d_{H\cdots O}$ distance increases, the HB critical point moves progressively away from the O–H negative Laplacian region and becomes surrounded by regions of positive Laplacian of increasing size and located within the H basin. Accordingly, the SF contributions from the H to the HB critical point become increasingly negative the larger the oxygen donor to oxygen acceptor separation.

Table 7.10 Source contributions at the hydrogen-bond CP along the reaction path for the approach of two water molecules and in a number of prototypical hydrogen-bonded complexes.[a,b,c]

$d_{D\cdots A}$, Å, system, hydrogen-bond class	ρ_b	$\nabla^2\rho_b$	S(H)%	S(H+D)%	S(H+D+A)%	S_{MolD}%
Reaction path for the approach of two water molecules						
3.25	0.010	0.041	−116.0	28.6	31.6	47.7
3.02, 1, IHB	0.016	0.067	−72.3	34.3	53.0	47.6
2.75	0.031	0.124	−35.5	38.7	71.0	48.4
2.50	0.056	0.216	−12.5	42.9	83.9	46.4
2.25	0.106	0.333	+2.8	42.5	88.7	46.2
2.00	0.199	0.208	+13.1	42.7	93.5	44.7
Prototypical hydrogen-bond complexes						
3.02, 1, IHB	0.016	0.067	−72.3	34.3	53.0	47.6
2.75, 2, PAHB	0.035	0.092	−14.4	38.7	69.7	–
2.54, 3, RAHB	0.056	0.148	+2.1	36.8	70.8	–
2.37, 3′, RAHB*	0.177	−0.425	32.2	40.7	89.5	–
2.43, 4, −(CAHB)	0.167	−0.392	32.1	40.4	90.3	43.8
2.41, 5, +(CAHB)	0.167	−0.415	31.4	41.0	92.7	43.5

[a] Ref. [67]; All quantities in a.u. if not otherwise stated.
[b] The hydrogen bond (HB) systems are labeled as in the text and in Fig. 7.8 and are classified according to the HB class they belong [71]. H, D, A are, respectively, the H directly involved in the HB, the hydrogen-donor and the hydrogen-acceptor O atoms. For system 2 only one HB is reported, the other two being very much alike. $S(\Omega', \Omega'', \ldots)$ expresses the sum of percentage source contributions from atoms $(\Omega', \Omega'', \ldots)$. S_{MolD}% is the percentage source contribution from the hydrogen-donor molecule, for those systems for which this molecule can be identified.
[c] System 3′ is classified as RAHB*, because it does not correspond to the energy minimum equilibrium configuration but to the TS for H atom migration.

At variance with the situation for typical covalent bonds, for which the sum of SF contributions from the two bonded atoms to their BCP ρ_b value is usually at least as large as 90% [67], the percentage contributions from the H and the donor oxygen, S(H+D)%, or from the H and the acceptor oxygen, S(H+A)%, are usually much smaller, if not even negative. This is the result of the hydrogen-atom behaving as a sink for most of the D···A distances and as a scarce source at very short distances. Percentage values close to 85–90% are obtained only when the sum of percentage contributions from all the three atoms directly involved in the HB, S(H+D+A)%, is considered, and only for $d_{H\cdots O}$ distances below 2.5 Å. This confirms the at least three-center nature of HBs [71]. As shown in Table 7.10, however, the sum of SF contributions from the remaining atoms in the molecular

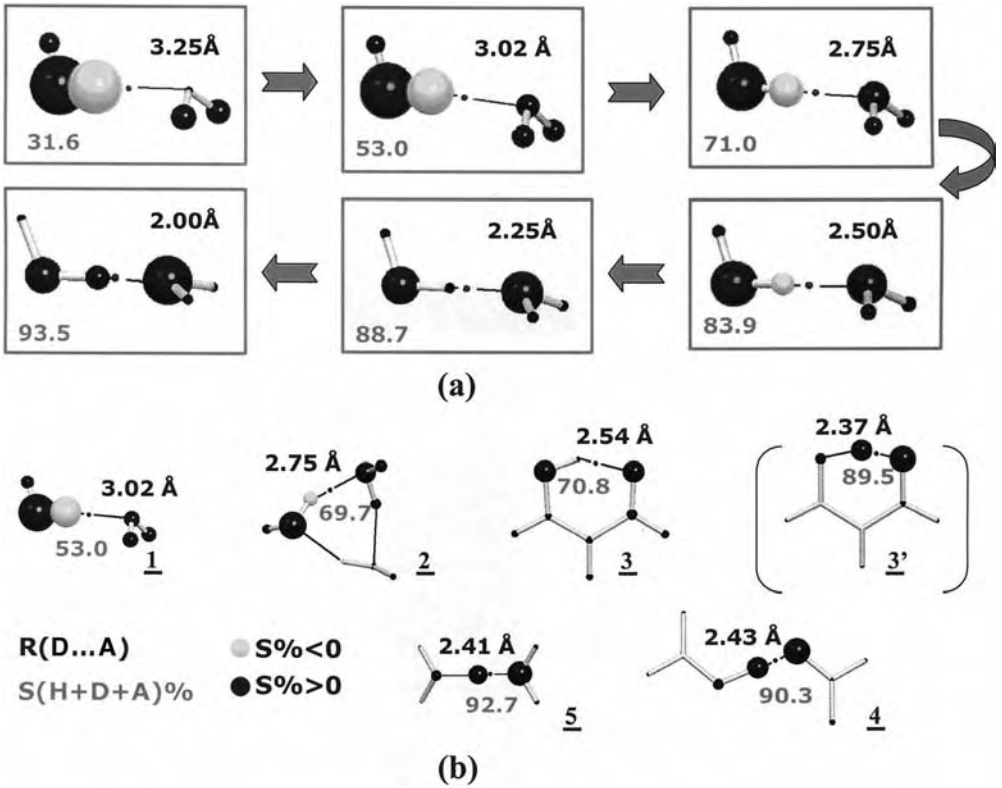


Fig. 7.8 Hydrogen-bonded molecular complexes – percentage atomic source contributions to the electron density at the HB critical point. (a) Changes of percentage atomic source contributions along the reaction path for the approach of two water molecules, within the linear C_s dimer constraint; (b) Typical hydrogen-bonded systems: 1 water dimer at equilibrium geometry; 2 cyclic homodromic water trimer; 3 malonaldeyde, C_s equilibrium form, and 3′, C_{2v} transition state for hydrogen migration; 4 open form of the formic acid–formate anion complex; 5 $H_5O_2^+$. Source contributions are displayed as balls whose volume is proportional to the percentage contribution from each atom, with positive sources dark shaded and negative sources light shaded. The black dot denotes the HB critical point for which atomic source contributions are evaluated. The source contribution from the H atom directly involved in the HB is a distinctive marker of the nature of the HB. S(H+D+A)% values are shown for each system and represent the sum of percentage sources from the H atom directly involved in the HB and from the hydrogen-donor and hydrogen-acceptor oxygen atoms. The sum of sources from the remaining atoms are as large as 47% in 1, an HB dominated by electrostatic interactions, and become less than 10% in the essentially covalent HBs 4 and 5. (Adapted, with permission, from Figs. 5 and 7 of Ref. [67], C. Gatti, F. Cargnoni, L. Bertini, *J. Comput. Chem.* 2003, *24*, 422–436, Copyright 2003, Wiley).

complex becomes large past 2.5 Å and, for instance, is as large as 47% at equilibrium. With the exception of short $d_{H\cdots O}$ distances, the description of the OH $\cdots$ O bond as a three-center interaction seems far from reality. As the donor-to-acceptor distance increases, other atoms in the complex become progressively involved in the hydrogen-bond (Fig. 7.8a), in accord with the increasing electrostatic character of the interaction.

The trend of SF contributions at the HB critical point occurring in real, prototypical hydrogen-bonded systems can now be easily appreciated. Following the Gilli and Gilli HB classification [71], Table 7.10 report data for:

1. the water dimer at equilibrium geometry, an isolated HB (IHB);
2. the cyclic homodromic water trimer, a polarization-assisted HB, (PAHB);
3. malonaldeyde, in its C_s equilibrium form, a resonance-assisted HB (RAHB);
4. the open form of the formic acid–formate anion complex, a negative charge-assisted HB, (−)CAHB; and
5. the symmetrical $[H_2O\cdots H\cdots OH_2]^+$ dimer, a positive charge assisted HB, (+)CAHB.

Also listed in the table are the data for the C_{2v} malonaldehyde TS (3′) for hydrogen-atom transfer between the two oxygen atoms. The percentage source from the H is corroborated as a characteristic, ideal marker of the nature of the hydrogen-bond, with changes in value as a function of the nature of HB that are even more pronounced than for the water dimer at different distances (Table 7.10 and Fig. 7.8b). This same behavior is observed for changes of percentage sources from other atoms and groups of atoms. The role of atoms other than the triad of atoms directly involved in the hydrogen-bond is clearly confirmed. The sum of the contributions from their sources is large for the IHB (47.0%), intermediate for the PAHB (30.3%) and the RAHB, and, indeed, small for the ±CAHB complexes, consistent with the increased HB covalence along the series. Interesting are the values for malonaldeyde, the RAHB model system. The percentage contribution from the H+D+A triad of atoms is as low as in the water trimer PAHB, which has a donor-to-acceptor distance approximately 0.2 Å larger. It seems that the RAHB molecule 3 is characterized by an enhanced source contribution from atoms other than the H+D+A triad and by local character of the D–H and D–A interactions less than expected on the basis of the donor–acceptor separation only. Similar features occur in the transition state 3′, which has more delocalized sources than the −CAHB system 4, despite a smaller $d_{D\cdots A}$ distance than 4. Further details on the classification of HBs according to the SF are reported in the original paper [67], with an interesting parallel between such a classification and that afforded [72] using the ELF.

We move now to the metal–metal bond. Since its discovery back in the 1960s, the metal–metal bond has kept challenging theory to enable detailed understanding of its nature in many transition metal molecular complexes. The advent and

increasing acceptance of topological analysis of the chemical bond have made it manifest that many traditional ideas of bonding are not applicable to these atypical bonds [73, 74]. Yet, even within the realm of topological approaches, well established bond classification schemes must be critically reconsidered and revised when applied to metal–metal bonds [21, 75]. For example, bond classification based on the sign of $\nabla^2\rho_b$ has proved to be largely inadequate [75] for bonds that usually involve atoms missing the outermost regions of charge depletion and concentration in their atomic Laplacian distributions and that have very low $|\nabla^2\rho_b|$ values, which makes the sign of $\nabla^2\rho_b$ quite indeterminate and the use of $\nabla^2\rho_b$, as the only classification index deceiving. Other quantities, for example the QTAIM delocalization indices, the $|V_b/G_b|$ ratio [76], the energy density H_b, the bond degree, H_b/ρ_b [76], the synaptic order, and the population of ELF valence basins [77, 78] have all been proposed as more informative topological indices in these instances [75, 79, 80]. Their application to an X-ray charge density study is in practice precluded, however, because evaluation of these indices requires the knowledge of the first (and second) density matrices which are, in general, not directly amenable to experiment. We have thus tested [69] SF analysis for several first transition row saturated and unsaturated binuclear homoleptic metal carbonyls, some of which have been seriously investigated both by theory and experiment [74, 81–83], and for the M_2(formamidinate)$_4$ complexes for which a comparative QTAIM and ELF study has already appeared [73].

The SF percentage contributions in $Mn_2(CO)_{10}$, a system with a formal metal–metal bond order of unity and whose bonding is still a matter of debate, are presented in Fig. 7.9 [74, 82, 83]. The Mn atoms act as a sink for the electron density

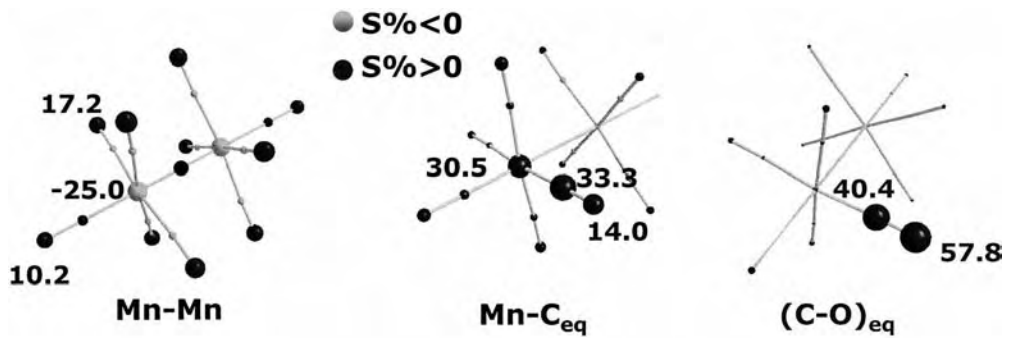


Fig. 7.9 Percentage atomic source contributions in $Mn_2(CO)_{10}$. Percentage sources at the Mn–Mn BCP (left), Mn–C_{eq} BCP (middle) and (C–O)$_{eq}$ BCP (right) are shown ($C_{eq} \equiv$ equatorial C). The percentage sources are portrayed with the same convention as in Fig. 7.8. Numerical values for the dominant positive and negative percentage sources are also displayed. Sources from the two atoms connected by the bond path are as high as 98.2% for the strong covalent–polar (C–O)$_{eq}$ bond, 63.8% for the Mn–C_{eq} dative bond, and highly negative (−50%) for Mn–Mn bond, where the carbonyl O atoms overbalance the electronic charge subtracted at BCP from the two Mn atoms.

at the metal–metal BCP and the carbonyl groups, essentially the carbonyl oxygen atoms, provide a positive contribution to ρ_b, so outweighing the negative source from the two Mn atoms. The SF views the Mn–Mn bond in $Mn_2(CO)_{10}$ as strongly nonlocalized, with the ligands inducing a noticeable polarization of the Laplacian distribution of the Mn basins so that they remove electronic charge from the bond midpoint. Conversely, a very different distribution of sources is observed for the dative $Mn–C_{eq}$ bonds and, in particular, for the (C–O)eq bonds, with the sum of percentage sources from the two linked atoms being 63.8% and 98.2%, respectively. Similar behavior is observed for the axial Mn–C and C–O bonds, with small changes in the percentage source values relative to the equatorial bonds, paralleling the corresponding small changes in the bond lengths. Although the contribution from Mn and C to the $Mn–C_{eq}$ BCP is positive and almost equally distributed between the two atoms, it is smaller than that of polar or nonpolar covalent bonds. The most important ''external'' contribution comes from the neighboring oxygen atom (14%), the other atoms in the molecule contributing the remaining 22.2%. For $(C–O)_{eq}$ the percentage sources comply with the supposedly covalent/polar nature of this bond, despite the positive, close to zero, Laplacian value found at the BCP. The polar character of the bond results in the O atom percentage source being significantly higher (57.8%) than that of the C atom (40.4%).

In the $Co_2(CO)_x$ ($x = 8–5$) series [84], source contributions from the Co atoms increase and become positive with decreasing net positive charge on the Co atom and with increasing formal Co–Co bond order, from one to four [69]. Trends of source contributions from Co atoms parallel the corresponding trends of δ(Co–Co) delocalization indices. Other proposed bond indices, such as $\nabla^2\rho_b$, the bond degree, H_b/ρ_b, and $|V|_b/G_b$, all fail to reproduce the trends in the δ(Co–Co) delocalization indices. In contrast, agreement between the source contributions and the delocalization indices persists even when the Co–Co BCP is lacking and the Co–Co mid-point is used to replace the Co–Co BCP as a reference point for evaluating the source contributions. The presence ($Co_2(CO)_8$, D_{3d}) or not ($Co_2(CO)_8$, C_{2v}) of a Co–Co BCP is, however, mirrored in characteristic different local source profiles along the Co–Co internuclear axis in the two cases. Although this analysis [69] has been performed on theoretically derived electron densities, it may be easily extended with no approximations to charge densities obtained from experiment, if available.

7.6.3
The Source Function: Latest Developments

The SF may be used to introduce an unambiguous full-electron population analysis [67]. By integrating Eq. (6) over a basin Ω, Eq. (8) is obtained:

$$N(\Omega) = \int_\Omega \rho(\mathbf{r})\,d\mathbf{r} = \int_\Omega S(\mathbf{r},\Omega)\,d\mathbf{r} + \sum_{\Omega' \neq \Omega} \int_\Omega S(\mathbf{r},\Omega')\,d\mathbf{r} \qquad (8)$$

and $N(\Omega)$ may thus be decomposed in an *inner contribution* $N_i(\Omega)$ and in a *outer contribution* $N_o(\Omega)$:

$$N(\Omega) = M(\Omega,\Omega) + \sum_{\Omega' \neq \Omega} M(\Omega,\Omega') = N_i(\Omega) + N_o(\Omega) \tag{9}$$

$M(\Omega,\Omega)$ represents the contribution to the electron population of Ω from its own basin, whereas $M(\Omega,\Omega')$ is the contribution to this same population from basin Ω'. The matrix M is not usually symmetric and defines a full population analysis based only on the observable $\nabla^2\rho$. Preliminary results, and the noteworthy numerical difficulties encountered in implementation of this unambiguous population analysis, have recently been presented [85] and will soon be published [86]. Interesting correlations of $M(\Omega,\Omega)$ and $M(\Omega,\Omega')$ with, respectively, the localization and the delocalization indices, have been observed in several diatomic molecules and along the ethane, ethylene, acetylene series. Such population analysis may potentially be applied, with no approximations, to the experimental electron densities.

Extension of the SF approach and of population analysis based on the SF to the α and β-components of the spin-polarized densities is planned for the near future.

References

1 C. Gatti, *TOPOND-98: An Electron Density Topological Program for Systems Periodic in N (N = 0–3) Dimensions*, User's Manual (CNR-ISTM, Milano, 1999).

2 C. Gatti, V. R. Saunders, C. Roetti, Crystal field effects on the topological properties of the electron density in molecular crystals: The case of urea, *J. Chem. Phys.* **1994**, *101*, 10686–10696.

3 C. Gatti, F. Cargnoni, Recent advances of the TOPOND program, pages 125–128, Proceedings III Convegno Nazionale di Informatica Chimica, Napoli, Italy, 27 February–1 March **1997**.

4 C. Katan, P. Rabiller, C. Lecomte, M. Guezo, V. Oison, M. Souhassou, Numerical computation of critical properties and atomic basins from three-dimensional grid electron densities, *J. Appl. Cryst.* **2003**, *36*, 65–73.

5 P. Rabiller, M. Souhassou, C. Katan, C. Gatti, C. Lecomte, Accuracy of topological analysis of gridded electron densities, *J. Phys. Chem. of Solids* **2004**, *65*, 1951–1955.

6 R. F. W. Bader, *Atoms in Molecules, A Quantum Theory*, International Series of Monographs on Chemistry, Oxford Science Publications, Oxford. UK **1990**, Vol. 22.

7 http://www.istm.cnr.it/~gatti/TOPOND.ppt

8 V. R. Saunders, R. Dovesi, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, C. M. Zicovich-Wilson, *CRYSTAL98, User's Manual*, University of Torino, Torino, **1998**.

9 C. Pisani, R. Dovesi, C. Roetti, *Hartree–Fock Ab Initio Treatment of Crystalline Systems*, Lecture Notes in Chemistry **1988**, Vol. 48, p. 150.

10 A. D. Becke, K. E. Edgecombe, A simple measure of electron localization in atomic and molecular systems, *J. Chem. Phys.* **1990**, *92*, 5397–5403.

11 Y. A. Abramov, On the Possibility of Kinetic Energy Density Evaluation

from the Experimental Electron-Density Distribution, *Acta Cryst.* **1997**, *A53*, 264–272.
12 X. Fradera, M. A. Austen, R. F. W. Bader, The Lewis Model and Beyond, *J. Phys. Chem* **1999**, *A103*, 304–314.
13 P. L. A. Popelier, A fast algorithm to compute atomic charges based on the topology of the electron density, *Theor. Chem. Acc.* **2001**, *105*, 393–399.
14 R. F. W. Bader, C. Gatti, A Green's function for the density, *Chem. Phys. Lett.* **1998**, *287*, 233–238.
15 P. L. A. Popelier, A robust algorithm to locate automatically all types of critical points in the charge density and its Laplacian, *Chem. Phys. Lett.* **1994**, *228*, 160–164.
16 M. Morse, S. S. Cairns, *Critical Point Theory in Global Analysis and Differential Geometry*, Academic Press, New York **1969**.
17 C. Gatti, P. Fantucci, G. Pacchioni, Charge density topological study of bonding in lithium clusters, *Theor. Chim. Acta* **1987**, *72*, 433–458.
18 V. Luaña, P. Mori-Sánchez, A. Costales, M. A. Blanco, A. Martín Pendás Non-nuclear maxima of the electron density on alkaline metals, *J. Chem. Phys.* **2003**, *119*, 6341–6350.
19 G. H. K. Madsen, P. Blaha, K. Schwarz, On the existence of non-nuclear maxima in simple metals, *J. Chem. Phys.* **2002**, *117*, 8030–8035.
20 T. S. Koritsanszky, P. Coppens, Chemical Applications of X-ray Charge-Density Analysis, *Chem. Rev.* **2001**, *101*, 1583–1627.
21 C. Gatti, Chemical bonding in crystals: new directions, *Z. Kristallogr.* **2005**, *220*, 399–457.
22 C. Pisani, Local techniques for the ab initio quantum-mechanical description of the chemical properties of crystalline materials, *J. Mol. Struct. (Theochem)* **2003**, *621*, 141–147.
23 R. Dovesi, B. Civalleri, R. Orlando, C. Roetti, V. R. Saunders, Ab initio Quantum Simulation in Solid State Chemistry, Rev. Comp. Chem. **2005**, *21*, 1–125.
24 A. Volkov, C. Gatti, Y. Abramov, P. Coppens, Evaluation of net atomic charges and atomic and molecular electrostatic moments through topological analysis of the experimental charge density, *Acta Cryst.* **2000**, *A56*, 252–258.
25 http://harker.chem.buffalo.edu/public/topxd/
26 P. Coppens, *X-Ray Charge Densities and Chemical Bonding*, IUCr Texts on Crystallography, International Union of Crystallography, Oxford University Press **1997**, vol. 4, p. 67.
27 A. Volkow, P. Macchi, L. J. Farrugia, C. Gatti, P. Mallinson, T. Richter, T. Koritsanszky, XD2006: A Computer Program Package for Multipole Refinement, Topological Analysis of Charge Densities and Evaluation of Intermolecular Energies from Experimental and Theoretical Structure Factors.
28 http://xd.chem.buffalo.edu/intro.html
29 P. Coppens, A. Volkov, The interplay between experiment and theory in charge-density analysis, *Acta Cryst.* **2004**, *A60*, 357–364.
30 C. Gatti, R. Bianchi, R. Destro, F. Merati, Experimental vs. theoretical topological properties of charge density distributions. An application to the L-alanine molecule studied by X-ray diffraction at 23 K, *J. Mol. Struct. (THEOCHEM)* **1992**, *255*, 409–433.
31 A. Volkov, Y. Abramov, P. Coppens, C. Gatti, On the origin of topological differences between experimental and theoretical crystal charge densities, *Acta Cryst.* **2000**, *A56*, 332–339.
32 B. Schiøtt, J. Overgaard, F. K. Larsen, B. B. Iversen, Testing Theory Beyond Molecular Structure: Electron Density Distributions of Complex Molecules, *Int. J. Quant. Chem.* **2004**, *96*, 23–31.
33 A. Volkov, P. Coppens, Critical examination of the radial functions in the Hansen–Coppens multipole model through topological analysis of primary and refined theoretical densities, *Acta Cryst.* **2001**, *A57*, 395–405.
34 C. Gatti, E. May, R. Destro, F. Cargnoni, Fundamental properties

and nature of CH···O interactions in crystals on the basis of experimental and theoretical charge densities. The case of 3,4-bis(dimethylamino)-3-cyclobutene-1,2-dione (DMACB) crystal, *J. Phys. Chem.* **2002**, *A106*, 2707–2720.

35 E. May, R. Destro, C. Gatti, The unexpected and large enhancement of the dipole moment in the 3,4-bis(dimethylamino)-3-cyclobutene-1,2-dione (DMACB) molecule upon crystallization: A new role of the intermolecular CH···O interactions, *J. Am. Chem. Soc.* **2001**, *123*, 12248–12254.

36 S. Swaminathan, B. M. Craven, R. K. McMullan, The crystal structure and Molecular Thermal Motion of Urea at 12, 60 and 123 K from Neutron Diffraction, *Acta Cryst.* 1984, *B40*, 300–306 and references therein.

37 H. Birkedal, D. Madsen, R. H. Mathiesen, K. Knudsen, H.-P. Weber, P. Pattison, D. Schwarzenbach, The charge density of urea from synchrotron diffraction data, *Acta Cryst.* **2004**, *A60*, 371–381.

38 R. F. W. Bader, H. Essén, The characterization of atomic interactions, *J. Chem. Phys.* **1984**, *80*, 1943–1960.

39 C. B. Duke, Semiconductor Surface Reconstruction: The Structural Chemistry of Two-Dimensional Surface Compounds, *Chem. Rev.* **1996**, *96*, 1237–1259.

40 I. K. Robinson, X-ray Crystallography of Surfaces and Interfaces, *Acta Cryst.* **1998**, *A54*, 772–778.

41 G. P. Srivastava, Theory of semiconductor surface reconstruction, *Rep. Prog. Phys.* **1997**, *60*, 561–613.

42 F. Cargnoni, C. Gatti, E. May, D. Narducci, Geometrical reconstructions and electronic relaxations of silicon surfaces. I. An electron density topological study of H-covered and clean Si(111)(1×1) surfaces, *J. Chem. Phys.* **2000**, *112*, 887–899.

43 F. Cargnoni, C. Gatti, Direct space analysis of the Si–Si bonding pattern in the π-bonded chain reconstructed Si(111)(2×1) surface, *Theor. Chem. Acc.* **2001**, *105*, 309–322.

44 H. N. Waltenburg, J. T. Yates Jr, Surface Chemistry of Silicon, *Chem. Rev.* **1995**, *95*, 1589–1673.

45 N. W. Ashcroft, N. D. Mermin, *Solid State Physics*. Int. Edn. Saunders College, Philadelphia **1976**, pag. 353.

46 A. Kurokawa, S. Ichimura, High purity ozone oxidation on hydrogen passivated silicon surfaces, *Appl. Surf. Sci.* **1996**, *101*, 436–439.

47 K. C. Pandey, New π-Bonded Chain Model for Si(111)(2×1) Surface, *Phys. Rev. Lett.* **1981**, *47*, 1913–1917.

48 F. J. Himpsel, P. M. Marcus, R. Tromp, I. P. Batra, M. R. Cook, F. Jona, H. Liu, Structure analysis of Si(111)2x1 with low-energy electron diffraction, *Phys. Rev.* **1984**, *B30*, 2257–2259.

49 J. M. Lehn, Supramolecular chemistry: from molecular information towards self-organization and complex matter, *Rep. Prog. Phys.* **2004**, *67*, 249–265.

50 N. J. Turro, Molecular structure as a blueprint for supramolecular structure chemistry in confined spaces, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10766–10770.

51 G. S. Nolas in *Thermoelectric Handbook, Macro to Nano* (Ed.: D. M. Rowe), CRC Taylor and Francis, Boca Raton, FL, USA, **2006**, Chapter 33.

52 V. I. Srdanov, G. D. Stucky, E. Lippman, G. Engelhardt, Evidence for an Antiferromagnetic Transition in a Zeolite-Supported Cubic Lattice of *F* centers, *Phys. Rev. Lett.* **1998**, *80*, 2449–2452.

53 P. Rogl in *Thermoelectric Handbook, Macro to Nano* (Ed.: D. M. Rowe), CRC Taylor and Francis, Boca Raton, FL, USA, **2006**, Chapter 32.

54 G. A. Slack in *CRC Handbook of Thermoelectrics*, (Ed.: D. M. Rowe), CRC, Boca Raton, Boca Raton, FL, USA, **1995**, Chapter 34.

55 N. P. Blake, D. Bryan, S. Latturner, L. Möllnitz, G. D. Stucky, H. Metiu, Structure and stability of the clathrates $Ba_8Ga_{16}Ge_{30}$, $Sr_8Ga_{16}Ge_{30}$,

$Ba_8Ga_{16}Si_{30}$ and $Ba_8In_{16}Sn_{30}$, *J. Chem. Phys.* **2001**, *114*, 10063–10074.

56 A. Bentien, A. E. C. Palmqvist, J. D. Bryan, S. Latturner, G. D. Stucky, L. Furenlid, B. B. Iversen, Experimental charge densities of semiconducting cage structures containing alkaline earth guest atoms, *Angew. Chem. Int. Ed.* **2000**, *39*, 3613–3616.

57 A. Bentien, E. Nishibori, S. Paschen, B. B. Iversen, Crystal structures, atomic vibration, and disorder of the type-I thermoelectric clathrates $Ba_8Ga_{16}Si_{30}$, $Ba_8Ga_{16}Ge_{30}$, $Ba_8In_{16}Ge_{30}$, and $Sr_8Ga_{16}Ge_{30}$, *Phys. Rev.* **2005**, *B71*, 144107.

58 N. P. Blake, S. Latturner, D. Bryan, G. D. Stucky, H. Metiu, Band structures and thermoelectric properties of the clathrates $Ba_8Ga_{16}Ge_{30}$, $Sr_8Ga_{16}Ge_{30}$, $Ba_8Ga_{16}Si_{30}$, and $Ba_8In_{16}Sn_{30}$, *J. Chem. Phys.* **2001**, *115*, 8060–8073.

59 C. Gatti, L. Bertini, N. P. Blake, B. B. Iversen, Guest–Framework Interaction in Type I Inorganic Clathrates with Promising Thermoelectric Properties: On the Ionic versus Neutral Nature of the Alkaline-Earth Metal Guest A in $A_8Ga_{16}Ge_{30}$ (A = Sr, Ba), *Chem. Eur. J.* **2003**, *9*, 4556–4568.

60 L. Bertini, F. Cargnoni, C. Gatti in *Thermoelectric Handbook, Macro to Nano* (Ed.: D. M. Rowe), CRC Taylor and Francis, Boca Raton, FL, USA, **2006**, Chapter 7.

61 V. I. Srdanov, K. Haug, H. Metiu, G. D. Stucky, Na_4^{3+} Clusters in sodium sodalite, *J. Phys. Chem.* **1992**, *96*, 9039–9043.

62 G. K. H. Madsen, C. Gatti, B. B. Iversen, L. Damjanovic, G. D. Stucky, V. I. Srdanov, *F* center in sodium electrosodalite as a physical manifestation of a non-nuclear attractor in the electron density, *Phys. Rev.* **1999**, *B59*, 12359–12369.

63 W. L. Cao, C. Gatti, P. J. MacDougall, R. F. W. Bader, On the presence of non-nuclear attractors in the charge distributions of Li and Na clusters, *Chem. Phys. Lett.* **1987**, *141*, 380–385.

64 B. B. Iversen, F. K. Larsen, M. Souhassou, M. Takata, Experimental evidence for the existence of non-nuclear maxima in the electron-density distribution of metallic beryllium. A comparative study of the maximum entropy method and the multipole refinement method, *Acta Cryst.* **1995**, *B51*, 580–591.

65 R. F. W. Bader, J. A. Platts, Characterization of an *F*-center in an alkali halide cluster, *J. Chem. Phys.* **1997**, *107*, 8545–8553.

66 G. Arfken, *Mathematical Methods for Physicists*, 3rd ed., Academic Press, Orlando, Florida **1985**, p. 897.

67 C. Gatti, F. Cargnoni, L. Bertini, Chemical Information from the Source Function, *J Comput. Chem.* **2003**, *24*, 422–436.

68 C. Gatti, L. Bertini, The local form of the source function as a fingerprint of strong and weak intra- and intermolecular interactions, *Acta Cryst.* **2004**, *A60*, 438–449.

69 C. Gatti, D. Lasi, Source Function description of metal–metal bonding in *d*-block organometallic compounds, *Faraday Discussion*, **2007**, DOI: 10.1039/b605404h.

70 G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, New York **1997**, p. 12.

71 G. Gilli, P. Gilli, Towards an unified hydrogen-bond theory, *J. Mol. Struct.* **2000**, *552*, 1–15.

72 F. Fuster, B. Silvi, Does the topological approach characterize the hydrogen bond?, *Theor. Chem. Acc.* **2000**, *104*, 13–21.

73 R. Llusar, A. Beltrán, J. Andrés, F. Fuster, B. Silvi, Topological analysis of Multiple Metal–Metal Bonds in Dimers of the M_2(Formamidinate)$_4$ Type with M = Nb, Mo, Tc, Ru, Rh, and Pd, *J. Phys. Chem.* **2001**, *105*, 9460–9466.

74 R. Ponec, G. Yuzhakov, M. R. Sundberg, Chemical structures from the analysis of domain-averaged Fermi holes. Nature of the Mn–Mn bond in bis(pentacarbonylmanganese) *J. Comput. Chem.* **2005**, *26*, 447–454.

75 P. Macchi, A. Sironi, Chemical bonding in transition metal carbonyl

clusters: complementary analysis of theoretical and experimental electron densities, *Coordination Chem. Rev.* **2003**, *238–239*, 383–412.

76 E. Espinosa, I. Alkorta, J. Elguero, E. Molins, From weak to strong interactions: A comprehensive analysis of the topological and energetic properties of the electron density distribution involving X–H···F–Y systems, *J. Chem. Phys.* **2002**, *117*, 5529–5542.

77 A. Savin, B. Silvi, F. Colonna, Topological analysis of the electron localization function applied to delocalized bonds, *Can. J. Chem.* **1996**, *74*, 1088–1096.

78 B. Silvi, The synaptic order: a key concept to understand multicenter bonding, *J. Mol. Struct.* **2002**, *614*, 3–10.

79 M. Kohout, F. R. Wagner, Y. Grin, Electron localization function for transition metal compounds, *Theor. Chem. Acc.* **2002**, *108*, 150–156.

80 G. Gervasio, R. Bianchi, D. Marabello, About the topological classification of the metal–metal bond, *Chem. Phys. Lett.* **2004**, *387*, 481–484.

81 Y. Xie, J. H. Jang, R. B. King, H. F. Schaefer III, Binuclear Homoleptic Manganese Carbonyls: $Mn_2(CO)_x$ (X0 10, 9, 8, 7), *Inorg. Chem.* **2003**, *42*, 5219–5230.

82 R. Bianchi, G. Gervasio, D. Marabello, Experimental electron density analysis of $Mn_2(CO)_{10}$: Metal–Metal and Metal–Ligand Bond Characterization, *Inorg. Chem.* **2000**, *39*, 2360–2366.

83 L. J. Farrugia, P. R. Mallinson, B. Stewart, Experimental charge density in the transition metal complex $Mn_2(CO)_{10}$: a comparative study, *Acta Cryst.* **2003**, *B59*, 234–247.

84 J. P. Kenny, R. B. King, H. F. Schaefer III, Cobalt–Cobalt multiple bonds in homoleptic carbonyls? $Co_2(CO)_x$ (x = 5–8) structures, energetics, and vibrational spectra, *Inorg. Chem.* **2001**, *40*, 900–911.

85 D. Lasi, C. Gatti, Unambigous electron population analysis from the electron density observable?, *European Charge Density Meeting-IV*, ECDM-IV, 26–29 January **2006**, Branderburg on the Havel, Germany.

86 D. Lasi and C. Gatti in preparation.

8
Topology and Properties of the Electron Density in Solids

Víctor Luaña, Miguel A. Blanco, Aurora Costales, Paula Mori-Sánchez, and Angel Martín Pendás

8.1
Introduction

Most chapters in this book explore the foundation and consequences of the quantum theory of atoms in molecules [1–17] (QTAIM) when applied to gas-phase molecules. Although most of the theory remains unaltered for solids, there are some significant differences that should be taken into account.

First, solids are an exciting challenge to physical bonding theories, because of their wide diversity of macroscopic and microscopic behavior. Insulators and electrical conductors; ionic, covalent, and molecular crystals; a rich variety of magnetic conductors; impurities and defects that can modify local and even bulk properties; significant differences between bulk and surface electronic properties; and a yet to be understood effect of grain shape and size on electronic and mechanical solid properties. The field offers a large collection of problems awaiting and deserving a careful look.

Topologically the main difference between solids and gas phase molecules is that atomic basins always have a finite size inside solids. As a consequence, the electron density of solids always has a rich collection of ring and cage critical points, in addition to the usual bond and nuclear points that dominate the topological description of molecules. Finite basins also imply well defined atomic radii in every geometrical direction. In this way the topology of the electron density provides a firm foundation for the important concept in the solid state theory of the atomic (ionic) radius, which played a prominent role in the early theories of phase stability. We will observe how the radius concept emerges again in topological analysis of the electron density as we rationalize the different topologies observed in families of compounds.

The large variability and mixture of types of bonding that may be found together in a given crystal makes the solid state a sort of paradise for complex topological behavior. This includes all kinds of nonstandard connectivities, for

example bonds to rings, or bond to bonds, that are stabilized in high-symmetry situations [18], nonnuclear attractors which are rather uncommon in isolated molecules, and topological polytypism, i.e. systems with the same atomic arrangement with electron densities with different topologies.

Solids also change volume and shape, and undergo electronic and structural phase transitions under the influence of external thermodynamic conditions, most notably pressure and temperature. Several metastable phases can, in fact, coexist for the same thermodynamic regime. The graphite, diamond, and buckminsterfullerene allotropes of carbon come to mind as simple and well known examples of this polymorphism. The availability of external variables that may be controlled at will introduces a wide scenario that lies outside the capabilities of experimental molecular physics. Pressure, the state variable that is most easily simulated, may be used to push atoms well past their equilibrium positions in the gas phase, so many profound questions that are still the subject of debate may find a natural answer here. One example is the role of unusual bond paths, for example those occurring among the anions in ionic crystals. These are intrinsic bond paths, surviving the large geometrical changes that are introduced on compression. Another is the effect of the virial of the nuclear forces in determining atomic energies, which is obscured in geometrically constrained isolated molecules.

The mechanics of quantum atoms in solids also provides a rigorous microscopic formulation of thermodynamic properties [19, 20]. The exhaustive partition of volume into atomic contributions, for example, enables us to immediately partition some important physical properties of crystals, for example the macroscopic compressibility. This generates fruitful chemical images of atoms being differentially compressed under application of an external pressure, in accordance with their intrinsic, atomic compressibility.

Last, but not least, we cannot forget that the electron density in crystals is an experimental observable. The applications of the quantum theory of atoms in molecules have enjoyed an important boost in the last decade, coupled with the availability of increasingly accurate experimental densities [21]. Many, if not most, experimentalists have now embraced the QTAIM to rationalize their data, leaving behind all the problems associated with use of density difference maps. The day that experiments provide equivalent or even more accurate densities than careful computation is approaching very quickly.

Gatti and coworkers embraced QTAIM methods when discussing the electron density of L-alanine [22, 23], although Zou's work in Bader's laboratory was, probably, the first attempt to derive specifically crystalline concepts, for example the Wigner–Seitz cell, from the electron density topology [24, 25]. Whereas approximately 15 years have passed since these seminal topological works in crystals, the field is still in its infancy, and whole terrains remain unexplored. For example, the very-high-pressure regime, in which large bonding changes are expected to occur, has not been appropriately simulated, and the always difficult simulation of temperature has simply been ignored. These difficult problems, together with many others, guarantee plenty of free room for future work.

It is not our purpose in this chapter to provide a thorough review of the QTAIM work performed for solid materials. There are already excellent accounts that cover that work (for example, the review by Koritzansky and Coppens [21] on experimental densities, Gatti's excellent paper on computational results [26], and Chapters 10–12 of this book). We will focus, in contrast, on some of the basic topological information obtained when studying solids that is not present at the isolated gas phase level. For the rest of this chapter we will assume the solid to be in a crystalline state, so space-group symmetry will be an essential aspect of the determination and analysis of electron density. Our discussion will also be limited to ideal bulk properties like most of the studies published so far. (See Chapters 7 and 9 of this book for authoritative discussions of surfaces).

8.2
The Electron Density Topology and the Atomic Basin Shape

Solid state theory assumes an ideal crystal model is formed as a translationally invariant repetition of a parallelepipedic unit cell. Such an ideal crystal has no borders, but the $+\infty$ and $-\infty$ limits along each crystallographic direction are connected to form a ring. Whereas a finite molecule inhabits S^3, the ordinary three-dimensional space, the ideal crystal belongs in R^3, the three-dimensional torus. As a consequence, the Poincaré–Hopf formula, that connects the number of critical points (CP) of any molecular scalar field, is substituted here by the Morse relationships [27]:

$$n, c \geq 1, \quad b, r \geq 3, \quad n - b + r - c = 0, \tag{1}$$

where n, b, r, and c denote the number of $(3,-3)$ –nuclear–, $(3,-1)$ –bond–, $(3,+1)$ –ring–, and $(3,+3)$ –cage–, CPs, respectively, that can be found in *each* crystal unit cell. Apart from this traditional (rank, signature) notation, the four types of nondegenerate CP can be distinguished in terms of the dimensionality of their attraction and repulsion basins, from 0D (points) to 3D (volumes), as summarized in Table 8.1.

Space group symmetry plays an essential part in the determination and analysis of the crystalline electron density. The gradient vector field must fulfill the point group symmetry within the unit cell; it must, therefore, be aligned with the symmetry axes, be contained within the symmetry planes, and be zero wherever an inversion point occurs. As a consequence, the local point group symmetry of some positions within the unit cell ensures the gradient vector field will be zero there. For example, the coincidence of three noncollinear symmetry axes, of an axis and a perpendicular plane, or the existence of a single inversion center will force that particular place to be a CP. Table 8.2 summarizes the crystallographic symmetries that guarantee the presence of a CP.

The points with special symmetry within the unit cell have been catalogued for the 230 space groups [28, 29] under the denomination of Wyckoff positions.

Table 8.1 Critical points (CP) of the electron density classified by rank, signature, and the dimensions of its attraction and repulsion basins (AB and RB, respectively). The attraction (repulsion) basin is defined as the geometrical place of all uphill gradient vector field lines ending in (starting from) the CP. Only the nondegenerate CPs, i.e. those with rank 3, should occur in an ordinary molecule and crystal, the appearance of a degenerate point would indicate structural instability.

(r, s)	AB	RB	QTAIM name	Description
$(3,-3)$	3D	0D	Nucleus (n)	Local maximum
$(3,-1)$	2D	1D	Bond (b)	First-order saddle
$(3,+1)$	1D	2D	Ring (r)	Second-order saddle
$(3,+3)$	0D	3D	Cage (c)	Local minimum
$(2,-2)$	2D	0D		Degenerated 2D maximum
$(2,0)$	1D	1D		2D saddle
$(2,+2)$	0D	2D		Degenerated 2D minimum
$(1,-1)$	1D	0D		Degenerated 1D maximum
$(1,+1)$	0D	1D		Degenerated 1D minimum
$(0,0)$	0D	0D		Inflection point

Close inspection of Table 8.2 reveals that the sites ensured to be a CP are nothing but those Wyckoff positions having their three crystallographic coordinates fixed by symmetry. It is also interesting to notice that all other special positions, though not guaranteeing the occurrence of a CP, strongly limit their possible location. One-parameter and two-parameter special positions may have a null gradient point at selected values of those parameters, though the actual point may coincide with a higher symmetry position. In that way, the search for some hard-to-find CPs can be bracketed.

Table 8.3 shows the electron-density topology of face-centered cubic (FCC) Al. The unit cell of this metallic phase contains four Al atoms at Wyckoff's 4*a* posi-

Table 8.2 Symmetry of fixed-point positions ensuring the presence of a critical point.

Triclinic	$C_i(\bar{1})$				
Monoclinic	$C_{2h}(2/m)$				
Orthorhombic	$D_2(222)$	$D_{2h}(mmm)$			
Tetragonal	$C_{4h}(4/m)$	$D_4(422)$	$D_{2d}(\bar{4}2m)$	$D_{4h}(4/mmm)$	
Trigonal	$C_{3i}(\bar{3})$	$D_3(32)$	$D_{3d}(\bar{3}m)$		
Hexagonal	$C_{3h}(\bar{6})$	$C_{6h}(6/m)$	$D_6(622)$	$D_{3h}(62m)$	$D_{6h}(6/mmm)$
Cubic	$T(23)$	$T_h(m3)$	$O(432)$	$T_d(\bar{4}3m)$	$O_h(m\bar{3}m)$

Table 8.3 Electron density topology of Al for the experimental FCC crystal structure at ambient conditions (space group $Fm\bar{3}m$, $a = 4.0495$ Å). The results correspond to a WIEN [30] fpLAPW calculation using the PBE96 [31] generalized gradient approach (Section 8.7).

Wyckoff		Pos.	CP		
4*a*	$m\bar{3}m$	(0, 0, 0)	Al	1.518110^{3}	
4*b*	$m\bar{3}m$	(1/2, 1/2, 1/2)	c_1	1.656710^{-2}	$+2.236810^{-2}$
8*c*	$\bar{4}3m$	(1/4, 1/4, 1/4)	c_2	2.797510^{-2}	$+8.466810^{-3}$
24*d*	$m.mm$	(0, 1/4, 1/4)	*b*	2.995410^{-2}	-1.212710^{-3}
32*f*	$.3m$	(x, x, x) $x = 0.28574$	*r*	2.808310^{-2}	$+7.211810^{-3}$

tions, 24 Al–Al bond CPs, 32 ring CPs, and two different types of cage point, one at the 4*b* octahedral site and a second at the 8*c* tetrahedral position. Four out of these five types of CP occur at positions completely fixed by the local symmetry. Only the ring CPs happen along the (x, x, x) line, with a single free parameter.

As happens in molecules, the crystal volume is divided, by means of zero flux surfaces, into quantum regions in which all quantum observables are well defined and all quantum laws can be locally applied [1, 7, 14–16, 32]. No gradient field line can cross those surfaces except the single 1D repulsion (attraction) line of the bond (ring) CP whose 2D attraction (repulsion) basin is the origin of the surface.

The smallest quantum region in a crystal is a primary bundle [33], formed by all the gradient vector field lines of the electron density going uphill from a given cage CP to a given nuclear CP. The bundle is topologically equivalent to a polyhedron, with vertices, edges and faces in numbers that fulfills Euler's invariant formula:

$$faces + vertices = edges + 2. \tag{2}$$

Apart from the nuclear and the cage points that originate the bundle, the vertices are bond and ring CPs. The faces correspond to the surface of the 2D attraction basins of the bond, and to the 2D repulsion basins of the ring CPs. The intersection between these surfaces gives rise to the edges of the bundles. The most basic topological structure of the crystal is that of its distinct primary bundles, and of their interconnections.

Primary bundles are, however, not reported in common topological analyses of the electron density. They are normally collected to introduce coarser partitions of space in which bigger zero-flux bounded regions are taken as the primary objects of study.

Union of all the primary bundles sharing the same maximum gives rise to an *atomic basin*. The interior of this object is nothing but the attraction basin of its

central n CP. All the other types of CP lie at the boundaries of the atomic basins. Atomic basins, like the primary bundles, are topologically equivalent to polyhedra. Faces are now made of the complete 2D attraction basin of a bond CP, which is contained in the face; edges correspond to the 1D attraction basin of a ring CP, also contained in the edge; the vertices are cage points.

Whereas atomic basins have been the main topological object in all QTAIM studies published so far, there is a second way of collecting the primary bundles that deserves some attention. The union of all the bundles that share the same minimum will produce the 3D repulsion basin of the cage point. The other types of CP lie again at the surface of this object, but the mapping is now reversed – n points are the vertices, 1D repulsion basins of the b points form the edges, and the faces correspond to the 2D repulsion basin of r points.

The FCC Al crystal has two different types of primary bundle; both are shown in Fig. 8.1. The first is formed by the gradient lines from a c_1 cage point to an Al atom and contains four r and four b CPs on its surface. The second type is formed from a c_2–Al pair, and contains three r points and three b points. Most of the bundles that we have found on crystals respond to this pattern of $2 + 2m$ vertices, where $m \geq 3$ is the same number of r as of b points.

The atomic basin of every Al atom is formed (Fig. 8.2) by collecting six Al–c_1 and eight Al–c_2 bundles to form an easily recognizable rhombic dodecahedron. In a similar way, six bundles form the octahedral repulsion basin of the c_1 CPs, and four bundles fill in the tetrahedral hole of a c_2 CP.

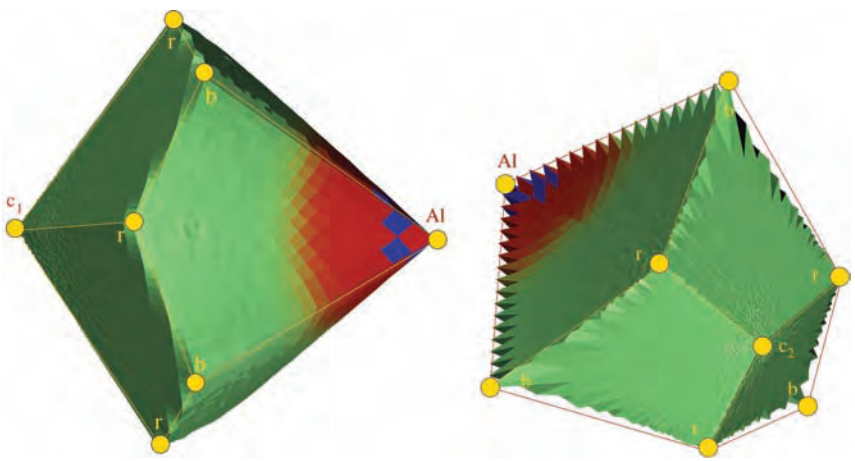


Fig. 8.1 Al–c1 and Al–c2 primary bundles for the FCC Al crystal. The bundle surfaces have been colored according to the $\nabla^2\rho$ scalar field to show the rapid variation of the electron density close to the Al nucleus. Notice that the bundles have n–b, b–r, and r–c edges, but not n–r nor b–c lines.

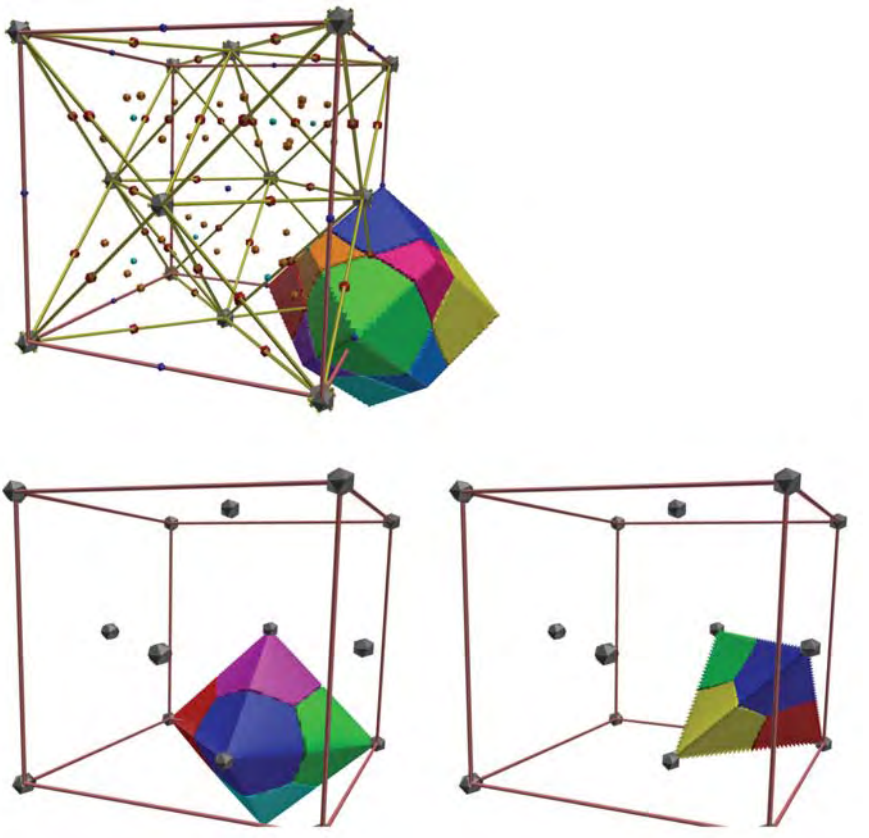


Fig. 8.2 Six Al–c_1 and eight Al–c_2 primary bundles form the atomic basin of Al. Four and six bundles, respectively, form the repulsion basin of c_1 and c_2 cage points. Each bundle is of a single solid color to show their matching in forming the coarser basin objects.

8.3
Crystalline Isostructural Families and Topological Polymorphism

A most significant fact emerges when the electron density is systematically analyzed in families of crystals – compounds with the same crystalline structure which are expected to have similar chemical properties have, however, different topologies of ρ. An example of this topological polytypism is provided by the alkaline-earth halides, MX_2, in the fluorite structure. These 20 compounds have one of three different topologies, represented in Fig. 8.3. What is most interesting is that the actual topology of a compound can be exactly predicted by determining the size ratio of the M and X species along the M–X bond path. When the M cation is less that 0.6 times the X anion, large faces created by the X–X bond critical

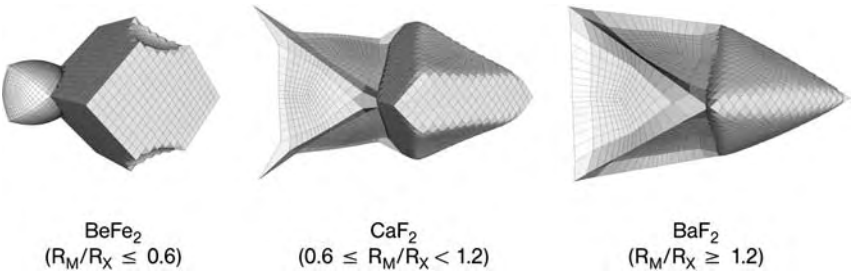


Fig. 8.3 Topological polytypism in alkaline-earth fluorites.

points dominate the anionic basin, which fills most of the crystal volume. The X–X bond paths continue to exist when R_M is in between 0.6 and 1.2 times R_X, but the size and effect of the X–X bonds decays progressively as we approach the upper limit. When $R_M/R_X \geq 1.2$, both M and X basins finally approach the shape dictated by their coordination in the crystalline structure (an octahedron and a tetrahedron, respectively) and the only bonding remaining is the M–X bond path between the nearest neighbors.

Similar geometrical reasoning explains the topological polytypism in other families we have examined, for example the B1 and B2 phases of alkali halides [33, 34], the alkaline/alkaline-earth perovskites AMX_3 [35], and their elpasolites A_2MX_4 [36]. The distance from a nucleus to its basin surface is very dependent on direction, but those directions determined by the bond paths contain the most useful information and enable us not only to systematize the occurrence of a given topology but also to explain trends in the thermodynamic properties of the crystals [37].

The examples above are of different compounds, albeit structurally related, with different topologies. A change of crystal structure as a result of a phase transition will also modify the electron density topology, but the interesting question is whether is it is possible for a single compound to have several different topologies without there being a structural change. We have found two different types of situation in which this behavior seems to be the norm.

The first example is the metallic phases of the low-electronegativity elements. BCC Li is the most extreme example, because it has approximately twenty different topologies for a change in the lattice size from 2.4 to 4.8 Å [18], but all the alkaline and alkaline-earth metals share the same tendency toward topological change. Perhaps most amazing is that the change follows a well defined trend, revealed when the many topologies are classified according to some significant features:

B_2: first and second metal neighbors are bonded;
B_1: only nearest neighbors bond CPs remain; and
M: the topology has nonnuclear maxima (NNM), also referred to as nonnuclear attractors (NNA), in addition to the prototypical nuclear $(3, -3)$ CPs [38–40].

The topology of the different metals follows a common trend on compression: $B_2 \rightarrow B_1[\rightarrow M] \rightarrow B2 \ldots$. The cause of this topological lability is that the electron density is almost constant, i.e. flat, throughout the valence region of these metals. An appropriate measurement of the *valence flatness* is provided by the ratio [41]:

$$f = \frac{\rho_c^{min}}{\rho_b^{max}} \tag{3}$$

where ρ_c^{min} is the absolute minimum of the electron density, and ρ_b^{max} is the maximum electron density found among the bond CPs. Typical covalent, ionic, and molecular solids have f values quite close to zero. Alkali metals, in contrast, approach $f = 100\%$ (98% Li, 95% Na and K, 91% Rb, and 88% Cs at their respective experimental geometries) and truly resemble the ideal Drude model of a free electron gas.

Topological polymorphism can also be expected for nonelementary compounds when the different elements share an extremely similar electronegativity. BP is a prototype for this example. Whereas the B3 zinc blende structure is maintained for a wide range of pressures and temperatures, we have found [42] that the standard polarity, $B^{+0.85}P^{-0.85}$ under ambient conditions, undergoes a reversal on application of hydrostatic pressure. The inversion occurs through an intermediate situation in which the P valence shell is transferred to a nonnuclear maximum before being caught by the B atom. In this way the low-pressure boron phosphide phase is transformed into high-pressure phosphorus boride, even though both are structurally equivalent.

8.4 Topological Classification of Crystals

The main topological properties of a prototypical set of compounds are listed in Table 8.4. A traditionally debated and often controversial question among solid state scientists is what constitutes the ionicity, covalence, and metallicity of a material. This question can be systematically answered with the aid of the magnitudes provided by the QTAIM analysis [41]. As is apparent from Table 8.4, the electron density at a bond CP, ρ_b, is characteristically large for covalent bonding. It is not easy, however, to compare the bond CP between two different pairs of elements in terms of theirs ρ_b values. The Laplacian of the electron density at a point, $\nabla^2\rho(\vec{r})$, measures whether the electron density is locally concentrated ($\nabla^2\rho < 0$) or depleted ($\nabla^2\rho > 0$) there, and provides a detailed map of the basic and acidic regions, respectively, of the crystal [43]. In a typically covalent bond, a region of negative Laplacian would include the bond path together with the two bonded nuclei. Prototypical ionic bonds, in contrast, would have spherical shells of basic character on each nucleus, the bond critical point occurring in a region of acidic character. Between both extreme examples we can find a collection of *polar bonding* situations with mixed signs. The ratio $-\lambda_1/\lambda_3$ enables measurement of

Table 8.4 Topological properties of a representative collection of crystals, including covalent, ionic, metallic, and molecular compounds. The values ρ_b and $\nabla^2\rho_b$ are the electron density and the Laplacian, respectively, at a bond CP. $\lambda_1 \leq \lambda_2 \leq \lambda_3$ are the eigenvalues of the Hessian, and the ratio $-\lambda_1/\lambda_3$ in the table corresponds to a bond CP. $Q(A)$ is the charge integrated within the basin of A. The last column is the valence density *flatness*, f (Eq. 3). The results correspond to HF-LCAO calculations (Section 8.7). All magnitudes are given in atomic units.

Crystal	*AB*	ρ_b	$\nabla^2\rho_b$	$-\lambda_1/\lambda_3$	*A*	$Q(A)$	f (%)
Diamond	C–C	0.2659	−0.9044	3.75	C	0.000	4.8
N_2	N–N	0.7380	−2.9228	4.13	N	0.000	0.0
	N–N	0.0023	0.0158	0.12			
	N–N	0.0012	0.0099	0.10			
CaF_2	Ca–F	0.0297	0.1780	0.15	Ca	1.821	1.3
	F–F	0.0127	0.0627	0.14	F	−0.911	
Li_2O	Li–O	0.0278	0.2055	0.15	Li	0.897	7.3
	O–O	0.0101	0.0317	0.17	O	−1.792	
Al	Al–Al	0.0308	0.0084	0.47	Al	0.000	56.8
Li	Li–NNM	0.0071	0.0067	0.02	Li	0.825	89.2
	NNM–NNM	0.0078	−0.0010	10.44	NNM	−0.137	

bond directionality, which is characteristically >1 for covalent bonds and <1 otherwise. Crystal ionicity, on the other hand, can be typically associated with large absolute values of the charge integrated on the basins. Metals are revealed by the flatness of the electron density in the valence region, which can now be measured quantitatively by the ratio of the electron density on the lowest density cage CP to the highest bond CP in the crystal. This flatness index, Eq. (3), can even be particularized to a plane or a line for low-dimensional metallic systems.

Among the crystals in Table 8.4, diamond and N_2 clearly have covalent bond CPs. For CaF_2 and Li_2O there is evidence of large ionicity, and Al and Li are characteristically metallic. For the N_2 crystal there is, in addition to the main N–N bond CP, another two types of bond path between neighboring molecules, which are clearly differentiated by their low density, positive and small Laplacian, and low bond directionality. For the Li crystal, in contrast, nonnuclear maxima of the electron density are observed. The NNM have been shown to occur in many homoatomic molecules and crystals when the appropriate internuclear distance regime is examined [38]. The appropriate distance range is, usually, much smaller than the typical equilibrium distances, and only a few atomic combinations can have NNM under normal thermodynamic conditions. These include metals, for example Li and Na, and also crystals, for example CaC_2. The occur-

rence of NNM is neither characteristic nor exclusive of metals, as was conjectured when this rare topological feature was first discovered.

8.5
Bond Properties – Continuity from the Molecular to the Crystalline Regime

One of the most important features of solid state behavior with regard to the application of QTAIM methods is enlargement of the narrow window of bonding ranges of molecules. In solids, external properties such as pressure and temperature, or the occurrence of polymorphism within a given substance, enable us to observe a given bond over a wide range of different distances within the same chemical environment, and even tune these distances at will. Whereas consideration of, for example, the N–N bond in N_2 at distances different from the equilibrium distance is theoretically easy, this is not achievable in the gas phase, and thus molecular quantum chemists seldom look into the dependence on distance of bond properties. This dependence has, in contrast, been the topic of several solid state studies both of ionic bonds [34, 37, 42, 44] and of intramolecular and intermolecular hydrogen bonds (by Espinosa et al. [45–47] and Gálvez et al. [48]). All these studies show that for the same bond (or type of bond) in different solid-state conditions there is clear dependence of bond densities (ρ_b) and Laplacians ($\nabla^2\rho_b$) on interatomic distance, usually linear in a logarithmic plot. Thus, geometric arguments related to ionic radii, for example, should be intimately related to bond properties, and these are, in turn, related to bond energetics [49]. These arguments can, however, be taken a step further; there is, in fact, a continuous bonding regime from the molecular to the crystalline limits [50, 51], with several universal features [52].

To unveil the relationship between the geometrical arguments and the different chemical behavior of bonds, it is most useful to seek approximate models that provide simplified, even analytical results. In this respect, the promolecular model (Section 8.7) provides an atomistic image of the chemical bond; using the sum of free atomic densities as an approximation, a good qualitative picture of the bond properties is obtained. This can be seen in Fig. 8.4, in which the main features of a high-level CISD N_2 Laplacian are reproduced qualitatively by the promolecular HF Laplacian. This dependence serves also to explain an important universal feature – the bond type, as given by $\nabla^2\rho_b$, is also a function of distance [52]. Thus, N_2, a covalently bonded molecule at the equilibrium distance (2.1 bohr) and in a certain range around it, has a positive Laplacian for distances larger than 3 bohr, and again for distances smaller than 1 bohr. This is already present in the simplified promolecular model, being related to a balance between the radial curvature and slope of the atomic densities. It is thus an atomic feature: the atomic shell structure is unchanged at large distances, shared at intermediate distances, where the shells interpenetrate substantially, and is completely fused at very short distances, where the K shells of the N atoms start to interact, initially

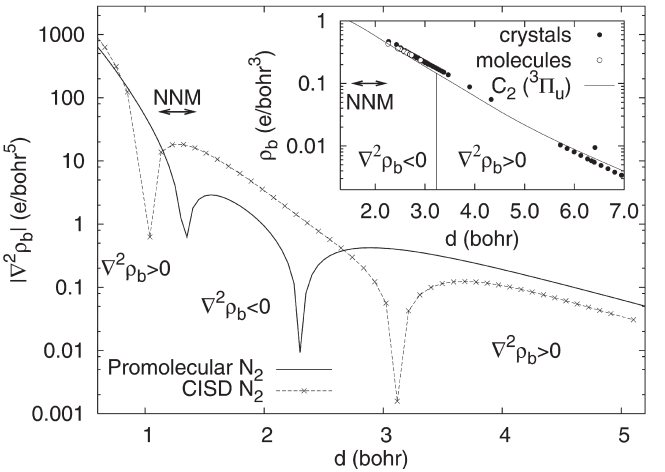


Fig. 8.4 Promolecular and CISD//TZV+(3d,1f) bond point Laplacians (absolute value, on a logarithmic scale) in the N_2 $X{-}^1\Sigma_g^+$ ground state. The inset shows C–C bond point densities for a set of B3LYP//6-311G(3df,p) molecules (C_2, ethane, ethene, ethine, benzene, anthracene, alene) and PBE-fpLAPW crystals (diamond and graphite at a wide range of distances, plus CaC_2), with the bond point density dependence on C–C distance for C_2 $^3\Pi_u$.

in a unshared manner. The same kind of behavior can be seen in other homodiatomics, for example Na_2 and Ne_2, with very different kinds of bond at the equilibrium distance, depending mainly on the range at which this distance falls for each of the species. The basic difference in heterodiatomics with quite different electronegativities (NaF and AlN) is that the valence shells do not fuse, but rather are totally or partially transferred from one atom to the other.

Further analysis of the parallel curvature in the bond point enables simple explanation of another important phenomenon – the existence of nonnuclear maxima of the electron density [18, 38]. There is, indeed, a window (marked in the graph) in which, for N_2, NNM are observed in the CISD density, and although this is not so for the promolecular density (its curvature falls short of becoming negative by a small amount), it is present for C_2 (inset of Fig. 8.4). It can thus be shown to be a feature present in many homonuclear bonds (and also some heteronuclear bonds [42]), and easily related to atomic shell structure.

In addition to this universal behavior of the bond density and Laplacian, with evolution quite similar in different compounds, they also have two more universal quantitative features. The first is apparent from analysis of the same type of bond in different molecules and crystals – the dependence of the bond properties on distance is the same for a given pair of atoms, continuous even for quite different types of bonding, as explained above. An example is seen in the inset of Fig. 8.4, in which data from single, double, and triple C–C bonds are depicted

with diamond and graphite crystal bond data for a wide range of distances. There is a general trend (slightly different for the long-distance interlayer van der Waals bonds in graphite, poorly described by these calculation) in all these situations which can be almost quantitatively fitted by the dependence of the behavior of the C_2 molecule on distance. Obviously, this is an even better model than the promolecule, because C–C bond features are included. Thus, the type of bonding is again a question of distance, which can also be read in the reverse way – the chemistry of the molecule determines the type of bond, and hence a given distance must occur. This interrelationship between geometric features and chemical behavior is behind the success of the geometric arguments previously given.

Finally, there is another kind of universal feature, by which bond properties can be transferred even across different bonding pairs with a common atom. To do this, the independent variable cannot be the interatomic distance, because this varies substantially with atomic sizes, but is instead the distance, r_A, from the common atom to the bond point. When this is done for a large number of group III nitride clusters, the corresponding B3-phase crystals, nitrogen hydrides, and N_2, the plot of bond density against r_N plot has a common trend, quite similar to the N_2 and promolecular trends [52]. To justify this trend, an exponential tail model of the atomic densities within the promolecular scheme, a further simplification, can be used to prove that, in fact, the homodiatomic and different heterodiatomic logarithmic plots differ approximately by a constant. This constant is not large, and depends on the different rates of decay of the atomic densities. Thus, the density (and to a lesser extent, its Laplacian) at the bond point is a universal function of the distance to a given atom, except for a small term that depends on the nature of the atom at the other end of the bond.

All of these arguments point in the same direction – the general features of bond properties mainly depend on geometry; because geometry depends, in turn, on bond type and strength, we can conclude that, as empirically proven previously, ρ_b and $\nabla^2\rho_b$ are good indicators of bond strength and type.

8.6
Basin Partition of the Thermodynamic Properties

In addition to a topological theory of chemical bonding, QTAIM is also a very successful atomic quantum theory – its definition of the atom as a proper open system, fulfilling atomic versions of the key theorems of quantum mechanics, provides a starting point for the following discussion. As is well known [49], QTAIM definition of atomic properties entails construction of appropriate local densities for the property at hand, and integrating these within the atomic basin, $O_A = \int_{\Omega_A} \rho_O(\vec{r})\, d\vec{r}$. This is termed an atomic average for the property, although the name is misleading – it is a sum of property O values at different points, not an average of values of the property. Then, for properties conforming with this, an atomic partition can be obtained, $O = \sum_A O_A$. Properties such as atomic

volumes (unit density), electron populations/atomic charges (electron density), kinetic energies (kinetic energy densities), ... can be evaluated in this way. (Chapter 1 of this book).

When considering a macroscopic crystal in the static limit, it is readily seen that extensive properties conform to the above scheme. Partitioning an extensive property within QTAIM is thus straightforward, by just defining the corresponding property density. Intensive properties, however, cannot be partitioned in this way. If any, a local associated value for the property can be defined instead of a density, and this local value should correspond to the macroscopic one via an appropriate average of the form $P = \sum_A w_A P_A$ [53]. The weight function w_a depends on the type of intensive property. If the derivative of an extensive property is involved (e.g. in the compressibility $\kappa = -V^{-1}(\partial V/\partial p)$), then $w_A = V_A/V$ [19] (and the local compressibility value will be $\kappa_A = -V_A^{-1}(\partial V_A/\partial p)$). If, on the other hand, derivatives with respect to an extensive property are involved (as in the bulk modulus $B = -V(\partial p/\partial V)$), it is found that $w_A = dV_A/dV$ (and the local bulk modulus is $B_A = -V_A(\partial p/\partial V_A)$. These sentences can be formalized by employing a linear operator technique [53], which goes beyond the purposes of this contribution.

An important point must be stressed here, however – this formalism, treating macroscopic properties as parameters for variation of local or global extensive properties (implicitly assuming a constant-p statistical ensemble in the κ and B determinations, through the $V(p)$ and $V_A(p)$ computed equations of state) is just one of various possibilities. For example, in the definition of pressure, one may use such a relationship to define $p = \sum_A w_A p_A$ with $p_A = -dE_A/dV_A$ and $w_A = dV_A/dV$ via the $E_A(V)$ and $V_A(V)$ implicitly defined functions, or one may use the local definitions of stress and pressure, as in Pendás 2002 [20]. The latter leads to a local partition factor $w_A = V_A/V$ and to an atomic pressure which is defined also as an average, $p_A = V_A^{-1} \int_{\Omega_A} p(\vec{r})\, d\vec{r}$. Both partitions lead to the same total value, but local values depend on how the ratio and the derivative weight factors are related (usually in a simple manner). Again, we will focus on the former scheme, with parametric $V(p)$ curves, because of its simplicity.

As an example of how this thermodynamic property partitioning works, we will first consider the AB_2O_4 spinel family ($A = \mathrm{Zn}, \mathrm{Mg}$; $B = \mathrm{Al}, \mathrm{Ga}$). Most oxide spinels have experimental B values very narrowly clustered around 200 GPa, a fact that requires theoretical justification. Instead of B, the simpler partition of the compressibility:

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right) = \sum_A \frac{V_A}{V} \times \left[-\frac{1}{V_A}\left(\frac{\partial V_A}{\partial p}\right)\right] = \sum_A w_A \kappa_A \tag{4}$$

will be used, so that:

$$\frac{1}{B} = \kappa = \sum_A w_A \kappa_A = \sum_A w_A \frac{1}{B_A} \tag{5}$$

Table 8.5 Local bulk moduli (B_Ω, in GPa) and volume fractions (w_Ω) for some AB_2O_4 spinels. Note that each w_Ω includes the multiplicity of Ω within the crystal unit cell.

Spinel	w_A	B_A	w_B	B_B	w_O	B_O
$MgAl_2O_4$	0.089	282	0.099	332	0.813	202
$MgGa_2O_4$	0.088	261	0.164	284	0.749	196
$ZnAl_2O_4$	0.136	246	0.095	335	0.769	203
$ZnGa_2O_4$	0.135	241	0.158	308	0.707	196

Now, by obtaining $V(p)$ from a series of *ai*PI calculations (Section 8.7), for a range of volumes, followed by fitting of an analytical equation of state, and doing the same for the QTAIM-obtained V_A values at each V (thus implicitly obtaining $V_A(p)$), we can evaluate local compressibilities, κ_A, and the weighting factors, $w_A = V_A/V$ [19]. This leads to the results gathered in Table 8.5.

From these results, the total B values are 216, 211, 216, and 214 GPa, with numerical fitting errors smaller than 1 GPa on each. Now, the reasons behind the $B \approx 200$ GPa constant value are clear from Eq. (5) – the larger contributions to the compressibility will be those with larger w_A and larger $1/B_A$, both cases pointing to the O atoms. For the volume fraction, not only is the O atom, anion-like, larger than any of the cation-like atoms, but also there are four of them per seven-atom formula unit and hence w_O contributes more than 70% in all crystals. For the local compressibility/inverse local bulk modulus, anion-like O atoms are more compressible than the hard, positively charged, *A* and *B* cation-like atoms, but the difference is not as large: B_O is approximately 200 GPa, divalent Mg and Zn have B_A values of approximately 270 and 245 GPa, and trivalent Al and Ga have B_B of approximately 335 and 290 GPa. Thus the average will be clearly dominated by the O atom, which has a remarkably similar compressibility in all four spinels considered, and is expected to behave in the same way in other spinels.

These arguments are general in their spirit but particular in their values. For example, the 20 alkali halides in the B1 phase [19] have quite different bulk moduli, ranging from 7.6 GPa in CsI to 80.9 GPa in LiF. The reason for this is mainly the large variation in sizes, with cations taking from less than 8% (LiI) to almost 70% (CsF) of the unit cell volume. There is also a large variability of local B_A, both for different atoms (36, 35, 20, 18, and 12 GPa for the cation in the *A*Cl series) and for the same atom in different compounds (31, 27, 14, 13, and 9 for the Cl atom in the same *A*Cl series). This is because of the much softer regions of the local $V_A(p)$ equation of state sampled by these compounds, compared with the hard spinels – divalent and trivalent cations are very small and incompressible, and the stronger electrostatic forces also place a large stress over the O atom, whereas the comparatively larger monovalent cations are much more com-

pressible which, together with the weaker electrostatics, results in less stress over the halide anions. Thus, partitioning of thermodynamic data into atomic components aids interpretation of their values and provides a powerful tool for qualitative predictions based on atomistic intuition.

8.7 Obtaining the Electron Density of Crystals

Most solid state electronic structure methods make use of pseudopotential techniques to avoid the explicit calculation of the atomic-like core regions. The valence electron density obtained from those calculations can, in principle, be completed by addition of the missing atomic core densities [54–56]. It must be noticed, however, that there are many different strategies for obtaining pseudopotentials, usually involving quite different partitions of the atomic density into core-like and pseudovalence contributions. An inappropriate match between the core and valence portions can result in spurious topological features. Whereas the field is very promising, some effort is still needed, in our opinion, to compare the results from the main types of pseudopotential currently in use. There is, therefore, a limited number of methods that provide the all electron crystalline electron density in and analytical form suitable for high-quality QTAIM studies.

Under the promolecular approximation the crystalline electron density is obtained as a sum of atomic contributions:

$$\rho(\vec{r}) = \sum_j \rho_j(\vec{r} - \vec{R}_j) \tag{6}$$

where the sum runs over all atoms, and $\rho_j(\vec{r} - \vec{R}_j)$ is the electron density of the atom situated at $\vec{R}_j$. This simple formula not only enables rapid approximate estimation but lies at the core of the experimental determination of the electron density from X-ray and neutron diffraction experiments. As a further simplification, the atomic densities can be assumed to retain the spherical symmetry of the free atoms. Although the promolecular model lacks the internuclear electron density accumulation characteristic of true covalent bonding, it has been observed to frequently retain the topological features of the density in nonmolecular crystals. Even such fine details as the tiny nonnuclear maxima of some alkaline metals [18] can be retained. In contrast, the promolecular model has been shown to miss intermolecular bond CPs, i.e. give a different topology, in some molecular crystal studies [57].

The ab initio perturbed ion (*ai*PI) [58, 59] method has been used quite successfully for calculation of the electron density of many ionic crystals [33–35, 37, 44, 51, 60]. The method solves the Hartree–Fock equations of a solid in a localized Fock space by partitioning the crystal wavefunction into local, weakly overlapping, group functions, each containing a fixed number of electrons and a single nucleus. The *ai*PI crystal density can be described as a promolecular-like formula

where the atomic contributions have been self consistently adapted to the crystal environment.

The CRYSTAL software of Dovesi et al. [61, 62] exploits the same MO-LCAO (crystal orbitals by linear combination of atomic Gaussian functions should be the motto here) approach that dominate quantum chemical studies in the molecular realm. Hartree–Fock (HF) and density functional theory (DFT) calculations have been used to examine the electron density topology of many ionic, covalent, molecular, and metallic crystals [26, 63–65]. It must be remarked, however, that use of the rich and flexible bases that are common within the molecular regime usually gives rise to linear dependencies and other technical problems that render the crystal calculation unsuitable. The molecular basis sets must then be truncated and Gaussian exponents less than 0.03–0.05 bohr^{-2}, approximately, become forbidden. To compensate for this problem, the lowest retained Gaussian exponents are usually massaged or optimized within the crystal. When those difficulties are taken into account, the method remains one of the two best techniques available to obtain the crystalline electron density. The next version of the software, release of which is planned for 2006, promises a definitive solution to the basis set problem and could produce a significant breakthrough in the field.

The other candidate for the best technique crown is the full potential linearized augmented plane wave (fpLAPW) formalism, implemented in the WIEN [30, 66] software by Blaha et al. The method uses different treatment for nonoverlapping spherical regions close to the nuclei, *muffin tins*, than for the internuclear regions. Atomic orbitals inside each muffin tin are described as the product of spherical harmonics and fully relativistic radial functions. Plane waves and a scalar relativistic treatment is applied to the internuclear region. The wave functions of both pieces are forced to match on value and slope at the muffin boundaries, but perfect agreement is not possible, because it would require inclusion of infinite spherical harmonics within each atomic sphere. Some care is then required to ensure that any remaining discontinuity in the density or its slope is not seen by the topological algorithms. Saturating the internuclear region with plane waves is, in contrast, both easy and inexpensive.

The fpLAPW, HF-LCAO, and *ai*PI topological properties of MgO and ZnO, two representative examples of the crystals that can be examined with the three techniques, are compared in Table 8.6. As usually occurs, the three techniques arrive at the same topology, which validates the very rapid and very stable *ai*PI procedure for analyzing trends and exploring the effect of geometry changes. They also agree on the properties of the main bond critical points. The *ai*PI calculation, however, underestimates the electron density at the ring and cage CPs and, in general, in the low-density regions of the crystal. The effect is more significant, because the average coordination index becomes smaller. The basin integrated charges provide clear evidence of this – the charges in the oxide basin agree for MgO (LCAO: -1.795, and *ai*PI: $-1.852e$) but differ substantially for ZnO (LCAO: -1.496, and *ai*PI: $-1.800e$). The difference between the fpLAPW and HF-LCAO results is mainly because of the correlation effects taken into account in the first calculation: DFT-LCAO results (not shown in the table) close the gap

Table 8.6 Comparison of the crystal topologies obtained by use of fpLAPW (PBE96 [31] GGA), HF-LCAO, and *ai*PI calculations on MgO (rock-salt structure, $a = 4.210$ Å) and ZnO (wurtzite structure, $a = 3.250$ and $c = 5.207$ Å). The (n,b,r,c) column provides the number of CPs of each type found in the unit cell. The r_A/r_B column gives the distance from the nuclei to a bond CP along the bond path. Atomic units are used throughout.

Crystal	Method	(n,b,r,c)	CP	$\rho(\vec{r}_c)$	$\nabla^2\rho$	r_A/r_B
MgO	fpLAPW	(8, 48, 48, 8)	b(Mg, O)	0.03911	0.21775	1.716/2.262
			b(O, O)	0.01770	0.04553	2.813/2.813
			r	0.01713	0.05651	
			c	0.00998	0.02402	
MgO	HF-LCAO	(8, 48, 48, 8)	b(Mg, O)	0.03599	0.25421	1.694/2.284
			b(O, O)	0.01611	0.04900	2.813/2.813
			r	0.01507	0.06309	
			c	0.00723	0.03096	
MgO	aiPI	(8, 48, 48, 8)	b(Mg, O)	0.03266	0.28543	1.704/2.274
			b(O, O)	0.01183	0.05363	2.813/2.813
			r	0.01075	0.06737	
			c	0.00444	0.02460	
ZnO	HF-LCAO	(4, 10, 8, 2)	b(Zn, O)	0.07870	0.47364	1.789/1.941
			b(Zn, O)	0.07554	0.44617	1.802/1.962
			b(Zn, O)	0.00650	0.02995	2.903/3.173
			r	0.00649	0.02990	
			r	0.00437	0.01926	
			c	0.00177	0.00850	
ZnO	aiPI	(4, 10, 8, 2)	b(Zn, O)	0.07961	0.51117	1.797/1.932
			b(Zn, O)	0.07568	0.48603	1.811/1.953
			b(Zn, O)	0.00363	0.02172	2.809/3.268
			r	0.00363	0.02171	
			r	0.00261	0.01354	
			c	0.00082	0.00466	

between the LCAO and LAPW techniques. The correlation tends to increase the electron density population of the bond CPs at the expense of the population of the low-density regions.

The theoretical electron densities can be corrected by using the crystal form factors, $F_{obs}(\vec{h})$, derived from experimental diffraction measurements [67, 68]. In principle, the electron density of the crystal can be obtained directly from the form factors:

$$\rho_{obs}(\vec{r}) = \sum_{\vec{h}} F_{obs}(\vec{h}) e^{-i2\pi\vec{h}\cdot\vec{r}} \tag{7}$$

where the sum runs over the infinite lattice planes. In practice, however, the number of form factors determined is severely limited by the experimental conditions. Equation (7) comprises a series of alternating terms and truncation of the sum to any finite number of terms, no matter how large, produces a density that does not satisfy the required analytical conditions. In particular, $\rho_{obs}(\vec{r})$ can have unphysical negative values and is full of spurious hills and valleys that render it not useful the gradient vector field.

The electron density obtained as a Fourier difference synthesis has much better analytical properties [71]:

$$\rho(\vec{r}) = \rho_m(\vec{r}) + \sum_{\vec{h}} \Delta F(\vec{h}) e^{-i2\pi\vec{h}\cdot\vec{r}} \tag{8}$$

where $\rho_m(\vec{r})$ is a smooth model density and

$$\Delta F(\vec{h}) = F_{obs}(\vec{h}) - \int\int\int \rho_m(\vec{r}) e^{i2\pi\vec{h}\cdot\vec{r}}\, d\vec{r} \tag{9}$$

is the difference between the values measured and calculated using ρ_m for plane $\vec{h}$. A particular form of promolecular model constitutes the basis for the Hansen–Coppens multipole refinement model [72], used on most current studies of the experimentally determined electron density [21], but nothing prevents use of CRYSTAL's LCAO or WIEN's fpLAPW as the source of the ρ_m. The theoretical electron density supplies the information missing on the incomplete set of experimental form factors and the experimental data correct any effect not accounted for in the theoretical calculation, for example imperfect treatment of electronic correlation or an incomplete basis set. This combination of theoretical and experimental data can be of great help when investigating problems that lie at the resolution limit of either approach.

This is illustrated for the Si crystal in Fig. 8.5, in which the electron density along a nearest-neighbors Si–Si line is plotted. Addition of purely atomic densities (i.e. the promolecular model) produces a bond CP at the $(0, 0, 0)$ position with a very low density, $0.05e$ bohr^{-3}. HF and BPW91-DFT LCAO calculations introduce nonnuclear attractors between nearest-neighbor silicon atoms that have been the subject of much speculation [73]. The HF result provides a high density, $0.098e$ bohr^{-3} at the Si–Si midpoint, which is slightly reduced to $0.090e$ bohr^{-3} at the correlated level. When these three model densities are corrected with the consolidated set of Cumming and Hart [69] and the Sake and Kato [70] set of form factors, all nonnuclear maxima disappear, and a well developed bond point appears. The corrected densities are, in fact, very similar, differing by less than $0.0015e$ bohr^{-3} in the bonding region.

The meaning of those results should be assessed with some care. The Si–Si distance in the RTP (room temperature and pressure) diamond phase lies at the upper limit of a window of existence of NNM [38]. This is one of the reasons silicon

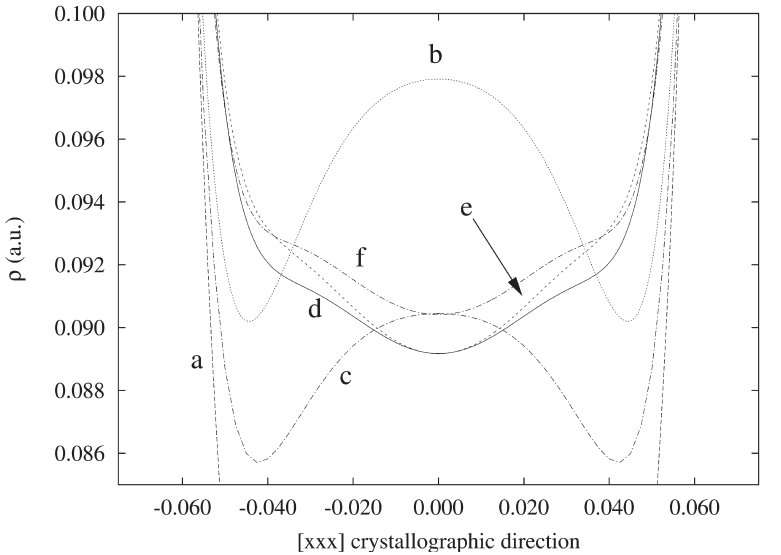


Fig. 8.5 Electron density of the Si crystal along the *xxx* crystallographic direction. This is the line connecting two nearest neighbors situated on $x = 1/8$ and $x = -1/8$, respectively. Lines (a) through (c) are the raw promolecular, HF-LCAO, and DFT-LCAO electron densities, respectively. Lines (d) through (f) correspond to the same model densities, now corrected with the experimental form factors [69, 70].

is an interesting target. The use of a small or unbalanced basis set, differences in the approximated treatment of the correlation, and other small computational details can modify the theoretical equilibrium distance enough to make the NNM appear or disappear. The experimental form factors are corrected, on their own, to avoid some well known effects, for example the directionally dependent X-ray absorption or the different time expended on measurement of the different $F_{obs}(\vec{h})$. The experimental form factors also unavoidably contain some thermal motion noise, even if they have been deconvoluted from thermal agitation. Neither quantum mechanical calculations nor experimental diffraction data are thus free from trouble when examining problematic systems lying at the edge of topological regions. The interesting thing is that both approaches can be combined in several ways, giving the researcher new tools to examine the stability of difficult topological features.

Acknowledgements

We thank the Spanish Ministerio de Ciencia y Tecnología for financial support under project BQU2003-06553.

References

1 R. F. W. Bader, Atoms in Molecules. A Quantum Theory, Oxford University Press, Oxford, 1990.
2 R. F. W. Bader, Int. J. Quantum Chem. 56 (1995) 409–419.
3 R. F. W. Bader, Phys. Rev. B 49 (1994) 13348–13356.
4 R. F. W. Bader, Int. J. Quantum Chem. 49 (1994) 299–308.
5 R. F. W. Bader, P. L. A. Popelier, Int. J. Quantum Chem. 45 (1993) 189–207.
6 P. F. Zou, R. F. W. Bader, Int. J. Quantum Chem. 43 (1992) 677–699.
7 R. F. W. Bader, Chem. Rev. 91 (1991) 893–928.
8 R. F. W. Bader, J. Chem. Phys. 91 (1989) 6989–7001.
9 R. F. W. Bader, Prog. Appl. Chem. 60 (1988) 145–155.
10 R. F. W. Bader, P. J. Macdougall, J. Am. Chem. Soc. 107 (1985) 6788–6795.
11 R. F. W. Bader, Acc. Chem. Res. 18 (1985) 9–15.
12 R. F. W. Bader, H. Essen, J. Chem. Phys. 80 (1984) 1943–1960.
13 R. F. W. Bader, T. T. Nguyen-Dang, Y. Tal, Rep. Prog. Phys. 44 (1981) 893–948.
14 R. F. W. Bader, T. T. Nguyen-Dang, Adv. Quantum Chem. 14 (1981) 63–124.
15 R. F. W. Bader, S. Srebrenik, T. T. Nguyen-Dang, J. Chem. Phys. 68 (1978) 3680–3691.
16 S. Srebrenik, R. F. W. Bader, T. T. Nguyen-Dang, J. Chem. Phys. 68 (1978) 3667–3679.
17 R. F. W. Bader, Acc. Chem. Res. 8 (1975) 34–40.
18 V. Luaña, P. Mori-Sánchez, A. Costales, M. A. Blanco, A. Martín Pendás, J. Chem. Phys. 119 (2003) 6341–6350.
19 A. M. Pendás, A. Costales, M. A. Blanco, J. M. Recio, V. Luaña, Phys. Rev. B. 62 (21) (2000) 13970–13978.
20 A. M. Pendás, J. Chem. Phys. 117 (2002) 965–979.
21 T. S. Koritsanszky, P. Coppens, Chem. Rev. 101 (2001) 1583–1627.
22 R. Destro, R. Bianchi, C. Gatti, F. Merati, Chem. Phys. Lett. 186 (1991) 47–52.
23 C. Gatti, R. Bianchi, R. Destro, F. Merati, J. Mol. Struct. (Theochem) 87 (1992) 409–433.
24 P. F. Zou, R. F. W. Bader, Acta Cryst. A 50 (1994) 714–725.
25 V. G. Tsirelson, P. F. Zou, T. H. Tang, R. F. W. Bader, Acta Cryst. A 51 (1995) 143–153.
26 C. Gatti, Z. Kristall, 220 (2005) 399–457.
27 M. Morse, S. S. Cairns, Critical point Theory in Global Analysis and Differential Geometry, Academic Press, New York, 1969.
28 T. Hahn (Ed.), International Tables for X-Ray Crystallography. A. Space-group symmetry, Vol. A of International Tables for X-Ray Crystallography, D. Reidel, Berlin, 1983.
29 M. I. Aroyo, J. M. Pérez-Mato, C. Capillas, E. Kroumova, S. Ivantchev, G. Madariaga, A. Kirov, H. Wondratschek. Zeitschrift für Kristallographie 221 (2006) 15–27, http://www.cryst.ehu.es/.
30 P. Blaha, K. Schwarz, J. Luitz, WIEN97, a Full Potential Linearized Augmented Plane Wave package for calculating crystal properties, (K. Schwartz, Techn. Univ. Wien, Vienna). Updated version of P. Blaha, K. Schwarz, P. Sorantin, and S. B. Trickey, Comput. Phys. Commun. **59**, 399, 1990 (1999).
31 J. P. Perdew, S. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865.
32 A. Martín Pendás, M. A. Blanco, E. Francisco, Chem. Phys. Letters 417 (2005) 16–21.
33 A. M. Pendás, A. Costales, V. Luaña, Phys. Rev. B 55 (1997) 4275–4284.
34 M. A. Blanco, A. Costales, A. M. Pendás, V. Luaña, Phys. Rev. B. 62 (2000) 12028–12039.

35 V. Luaña, A. Costales, A. Martín Pendás, Phys. Rev. B 55 (1997) 4285–4297.
36 A. Costales, Topología de la densidad electrónica en cristales. una teoría cuántica del enlace cristalino., Ph.D. thesis, Universidad de Oviedo (1998). URL http://web.uniovi.es/qcg/
37 A. Martín Pendás, A. Costales, V. Luaña, J. Phys. Chem. B 102 (1998) 6937–6948.
38 A. Martín Pendás, M. A. Blanco, A. Costales, P. Mori-Sánchez, V. Luaña, Phys. Rev. Lett. 83 (1999) 1930–1933.
39 G. K. H. Madsen, P. Blaha, K. Schwarz, J. Chem. Phys. 117 (2002) 8030–8035.
40 J. Friis, G. K. H. Madsen, F. K. Larsen, B. Jiang, K. Marthinsen, R. Holmestad, J. Chem. Phys. 119 (2003) 11359–11366.
41 P. Mori-Sánchez, A. Martín Pendás, V. Luaña, J. Am. Chem. Soc. 124 (2002) 14721–14723.
42 P. Mori-Sánchez, A. Martín Pendás, V. Luaña, Phys. Rev. B 63 (2001) 125103–1–4.
43 R. F. W. Bader, P. J. Macdougall, C. D. H. Lau, J. Am. Chem. Soc. 106 (1984) 1594–1605.
44 E. Francisco, J. M. Recio, M. A. Blanco, A. Martín Pendás, A. Costales, J. Phys. Chem. A 102 (1998) 1595–1601.
45 E. Espinosa, E. Molins, C. Lecomte, Chem. Phys. Lett. 285 (1998) 170–173.
46 E. Espinosa, M. Souhassou, H. Lachekar, C. Lecomte, Acta Cryst. B 55 (1999) 563–572.
47 E. Espinosa, I. Alkorta, J. Elguero, E. Molins, J. Chem. Phys. 117 (2002) 5529–5542.
48 O. Gálvez, P. C. Gómez, L. F. Pacios, J. Chem. Phys. 115 (2001) 11166–11184.
49 R. F. W. Bader, Atoms in Molecules. A Quantum Theory, Oxford University Press, Oxford, 1990.
50 A. Costales, M. A. Blanco, A. M. Pendás, A. K. Kandalam, R. Pandey, J. Am. Chem. Soc. 124 (2002) 4116–23.
51 V. Luaña, A. Costales, P. Mori-Sánchez, A. Martín Pendás, J. Phys. Chem. B 107 (2003) 4912–4921.
52 A. Costales, M. A. Blanco, P. Mori-Sánchez, V. Luaña, A. Martín Pendás, J. Phys. Chem. A 108 (2004) 2794–2801.
53 P. Mori-Sánchez, Densidad electrónica y enlace químico. De la molécula al cristal, Ph.D. thesis, Universidad de Oviedo (2002). URL http://web.uniovi.es/qcg/
54 J. Cioslowski, P. Piskorz, Chem. Phys. Lett. 255 (1996) 315–319.
55 S. F. Vyboishchikov, A. Sierraalta, G. Frenking, J. Comput. Chem. 18 (1997) 416–429.
56 A. Sierraalta, A. Herize, J. Mol. Struct. (Theochem) 529 (2000) 173–182.
57 C. Gatti, E. May, R. Destro, F. Cargnoni, J. Phys. Chem. A 106 (2002) 2707–2720.
58 V. Luaña, L. Pueyo, Phys. Rev. B 41 (1990) 3800–3814.
59 M. A. Blanco, V. Luaña, A. M. Pendás, Comput. Phys. Commun. 103 (1997) 287–302.
60 V. Luaña, A. Costales, A. M. Pendás, L. Pueyo, J. Phys.: Condens. Matter 11 (1999) 6329–6336.
61 R. Dovesi, V. R. Saunders, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, C. M. Zicovich-Wilson, CRYSTAL98 User's Manual (1998).
62 V. R. Saunders, R. Dovesi, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, N. M. Harrison, K. Doll, B. Civalleri, I. J. Bush, CRYSTAL03 User's Manual, http://www.crystal.unito.it/ (2003).
63 G. K. H. Madsen, C. Gatti, B. B. Iversen, L. Damjanovic, G. D. Stucky, V. I. Srdanov, Phys. Rev. B 59 (1999) 12359–12369.
64 C. Gatti, V. R. Saunders, C. Roetti, J. Chem. Phys. 101 (1994) 10686–10696.
65 B. Silvi, C. Gatti, J. Phys. Chem. A 104 (2000) 947–953.
66 P. Blaha, K. Schwarz, G. Madsen, D. Kvasnicka, J. Luitz, WIEN2k, an Augmented Plane Wave plus Local Orbitals program for calculating crystal properties, Vienna University of Technology, Institute of Physical

and Theoretical Chemistry (2001). URL http://www.wien2k.at/
67 G. H. Stout, L. H. Jensen, X-ray structure determination: A Practical Guide, 2nd Edition, John Wiley & Sons, New York, 1989.
68 C. Giacovazzo, H. L. Monaco, D. Viterbo, F. Scordari, G. Gilli, G. Zanotti, M. Catti, Fundamentals of Crystallography, Oxford UP, Oxford, UK, 1992.
69 S. Cumming, M. Hart, Aust. J. Phys. 41 (1988) 423.
70 T. Saka, N. Kato, Acta Crystallogr. A 42 (1986) 469.
71 R. Hosemann, S. N. Bagchi, Nature 171 (1953) 785–787.
72 N. K. Hansen, P. Coppens, Acta Cryst. A 34 (1978) 909.
73 J. M. Zuo, P. Blaha, K. Schwarz, J. Phys. Condens. Matter. 9 (1997) 7541–7561.

9
Atoms in Molecules Theory for Exploring the Nature of the Active Sites on Surfaces

Yosslen Aray, Jesus Rodríguez, and David Vega

9.1
Introduction

The atoms in molecules theory provides a rigorous definition of chemical bonds for all types of molecules and solids [1–10] and has proven useful in analysis of the physical properties of insulators, pure metals, and alloys [4–8]. In particular, it has been observed that the strength of the bonding between a given pair of atoms in a molecule correlates with the values of the electron density at the *bond critical point*, ρ_b [1]. A simple relationship between the binding energy and ρ_b in periodic solids has also been reported [4, 6, 8]. The AIM theory also leads to unique partition of three-dimensional space into a collection of chemically identifiable regions called atomic basins (the atoms in a molecule or in a crystal). These are the most transferable pieces one can define in exhaustive partitioning of the real space [1]. In this chapter we describe the implementation of an algorithm that uses the $\rho(\mathrm{r})$ topological information to determine the main elements of the AIM theory for periodic systems, and discuss an application to nanocatalysis.

9.2
Implementing the Determination of the Topological Properties of $\rho(\mathbf{r})$ from a Three-dimensional Grid

The CPs are usually calculated using the Newton–Raphson (NR) technique [11]. The NR algorithm starts from a truncated Taylor expansion at a point $\mathbf{r} = \mathbf{r}_0 + \mathrm{h}$, about $\mathbf{r}_0$ of a multidimensional scalar function ($\nabla\rho(\mathbf{r})$):

$$\nabla\rho(\mathbf{r}) = \nabla\rho(\mathbf{r}_0) + \mathbf{H}_0\mathbf{h} + \text{higher order terms} \tag{1}$$

where $\mathbf{H}$ is the Hessian (the Jacobian of $\nabla\rho(\mathbf{r})$) at point $\mathbf{r}_0$. The best step, $\mathbf{h}$, to move from the initial $\mathbf{r}_0$ to the CP is $\mathbf{h} = -\mathbf{H}^{-1}\nabla\rho(\mathbf{r}_0)$. This correction is then used to obtain a vector $\mathbf{r}_{\text{new}} = \mathbf{r}_{\text{old}} + t\mathbf{h}$ (t is a small value) and the process is iter-

ated to $\nabla\rho(\mathbf{r}) = 0$. The NR algorithm requires evaluation of the first and second partial derivatives of ρ, at arbitrary points **r**.

The points of the necessary gradient paths to determine the bond paths, crystal graphs, and IAS are solutions of the differential equation [1]:

$$d\mathbf{r}(s)/ds = \nabla\rho(\mathbf{r}(s)) \tag{2}$$

where the notation $\mathbf{r}(s)$ implies that a point **r** on a given path is dependent upon the path parameter s. Equation (2) represents three first-order differential equations ($dx_i(s)/ds = \partial\rho/\partial x_i$, $x_i = x, y, z$) and yields unique solutions only when particular values are assigned to three constants of integration. This corresponds to fixing some initial point on a trajectory, at $s = s1$, for example. A trajectory of the gradient vector field of $\rho(\mathbf{r})$ is a parametrized integral curve, a solution curve, of the differential equation for $\nabla\rho(\mathbf{r})$. By fixing a point on a given trajectory all other points which lie on the same path, can be obtained by solving Eq. (2). This is achieved by using a fifth-order Cash–Karp Runge–Kutta (CKRK) method [12]. The general form of the Runge–Kutta formula is:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + c_1\mathbf{k}_1 + c_2\mathbf{k}_2 + c_3\mathbf{k}_3 + c_4\mathbf{k}_4 + c_5\mathbf{k}_5 + c_6\mathbf{k}_6 \tag{3}$$

where $\mathbf{x}_n = (x_n, y_n, z_n)$ and $\mathbf{k}_j$ over an interval h are:

$$\mathbf{k}_1 = h\nabla\rho(\mathbf{r})|_{\mathbf{r}=\mathbf{x}_n}$$

$$\mathbf{k}_2 = h\nabla\rho(\mathbf{r})|_{\mathbf{r}=\mathbf{x}_n+b_{21}\mathbf{k}_1}$$

$$\ldots$$

$$\mathbf{k}_6 = h\nabla\rho(\mathbf{r})|_{\mathbf{r}=\mathbf{x}_n+b_{61}\mathbf{k}_1+\cdots+b_{65}\mathbf{k}_5}$$

The particular values of the various constants (c_j, b_{ij}) are given in Ref. [12].

It is apparent the key for implementation of the NR and CKRK algorithms is calculation of the required derivatives of $\rho(\mathbf{r})$, at arbitrary points **r**. To develop a method enabling us to study complex systems, irrespective of the basis set (analytically or numerically) used, a numerical method on electron densities given on regular, not necessarily homogeneous three-dimensional grids was implemented. The necessary partial derivatives are evaluated using a five-degree Lagrange polynomial interpolation of $\rho(\mathbf{r})$ and are fed into an automated algorithm for systematic determination of the all CPs. For just one dimension the interpolating polynomial of degree $n-1$ through n points $y_1 = f(x_1)$, $y_1 = f(x_1), \ldots, y_n = f(x_n)$ is given by the Lagrange's formula:

$$P(x) = \sum_{k=1}^{n} \frac{\prod_{j=1, j\neq k}^{n} (x - x_j)}{\prod_{j=1, j\neq k}^{n} (x_k - x_j)} y_k \tag{4}$$

There are n terms, each a polynomial of degree $n-1$ and each constructed to be zero at all x_j except one, at which it is constructed to be y_k. For a homogeneous grid, $x_j = x_1 + (j-1)\mathrm{h}$, where h is the step size. Defining $s = (x - x_\alpha)/h$ so that x_α and $x_{\alpha+1}$ are the central points of the grid, we obtain $x = x_\alpha + sh$. Substituting this last expression in Eq. (4) we have:

$$P = \sum_{k=1}^{n} \frac{\prod_{j=1,\, j\neq k}^{n} (\alpha - j + s)}{\prod_{j=1,\, j\neq k}^{n} (k-j)} y_k \tag{5}$$

and

$$P = \sum_{k=1}^{n} w_{k,n}(s) \cdot y_k \tag{6a}$$

where

$$w_{k,n}(s) = \frac{\prod_{j=1,\, j\neq k}^{n} (\alpha - j + s)}{\prod_{j=1,\, j\neq k}^{n} (k-j)} \tag{6b}$$

$w_{k,n}(s)$ are polynomials of degree $n-1$ in s. The v-degree derivative of these expressions is:

$$\begin{aligned} \frac{d^v P}{dx^v} &= \frac{1}{h^v} \sum_{k=1}^{n} w_{k,n}^{(v)}(s) \cdot y_k \\ w_{k,n}^{(v)}(s) &= \frac{d^v w_{k,n}(s)}{ds^v} \end{aligned} \tag{7}$$

Equations (6) and (7) provide an accurate, rapid, and efficient way of interpolating $\rho(\mathbf{r})$ and its derivatives at many arbitrary points. The easiest way to determine those derivatives is to develop $w_{k,n}(s)$ as an s polynomial. For example, in Table 9.1 the s polynomial expression for $w_{k,n}(s)$ and its first and second derivatives for interpolation with $n = 4, 6$, and 8 points are reported. The explicit expression for $n = 4$ (Fig. 9.1) is given by:

$$\begin{aligned} \rho(\mathrm{x}) = {} & (-s^3 + 3s^2 - 2s)/6\rho_1 + (s^3 - 2s^2 - s + 2)/2\rho_2 \\ & + (-s^3 + s^2 + 2s)/2\rho_3 + (s^3 - s)/6\rho_4 \end{aligned} \tag{8}$$

For a three-dimensional system P is evaluated at points surrounding a box containing the current point and in such a way that any coordinates of the point

Table 9.1 $w_{k,n}^{(v)}(s)$ polynomial defined using Eqs (6) and (7), with $\alpha = n/2$ and $w_{k,n}(s) = w_{k,n}^{(0)}(s)$.

		$n = 4$	$n = 6$	$n = 8$
$v = 0$	$k = 1$	$(-s^3 + 3s^2 - 2s)/6$	$(-s^5 + 5s^4 - 5s^3 - 5s^2 + 6s)/120$	$(-s^7 + 7s^6 - 7s^5 - 35s^4 + 56s^3 + 28s^2 - 48s)/5040$
	$k = 2$	$(s^3 - 2s^2 - s + 2)/2$	$(s^5 - 4s^4 - s^3 + 16s^2 - 12s)/24$	$(s^7 - 6s^6 - 2s^5 + 60s^4 - 71s^3 - 54s^2 + 72s)/720$
	$k = 3$	$(-s^3 + s^2 + 2s)/2$	$(-s^5 + 3s^4 + 5s^3 - 15s^2 - 4s + 12)/12$	$(-s^7 + 5s^6 + 9s^5 - 65s^4 + 16s^3 + 180s^2 - 144s)/240$
	$k = 4$	$(s^3 - s)/6$	$(s^5 - 2s^4 - 7s^3 + 8s^2 + 12s)/12$	$(s^7 - 4s^6 - 14s^5 + 56s^4 + 49s^3 - 196s^2 - 36s + 144)/144$
	$k = 5$		$(-s^5 + s^4 + 7s^3 - s^2 - 6s)/24$	$(-s^7 + 3s^6 + 17s^5 - 39s^4 - 88s^3 + 108s^2 + 144s)/144$
	$k = 6$		$(s^5 - 5s^3 + 4s)/120$	$(s^7 - 2s^6 - 18s^5 + 20s^4 + 89s^3 - 18s^2 - 72s)/240$
	$k = 7$			$(-s^7 + s^6 + 17s^5 - 5s^4 - 64s^3 + 4s^2 + 48s)/720$
	$k = 8$			$(s^7 - 14s^5 + 49s^3 - 36s)/5040$
$v = 1$	$k = 1$	$(-3s^2 + 6s - 2)/6$	$(-5s^4 + 20s^3 - 15s^2 - 10s + 6)/120$	$(-7s^6 + 42s^5 - 35s^4 - 140s^3 + 168s^2 + 56s - 48)/5040$
	$k = 2$	$(3s^2 - 4s - 1)/2$	$(5s^4 - 16s^3 - 3s^2 + 32s - 12)/24$	$(7s^6 - 36s^5 - 10s^4 + 240s^3 - 213s^2 - 108s + 72)/720$
	$k = 3$	$(-3s^2 + 2s + 2)/2$	$(-5s^4 + 12s^3 + 15s^2 - 30s - 4)/12$	$(-7s^6 + 30s^5 + 45s^4 - 260s^3 + 48s^2 + 360s - 144)/240$
	$k = 4$	$(3s^2 - 1)/6$	$(5s^4 - 8s^3 - 21s^2 + 16s + 12)/12$	$(7s^6 - 24s^5 - 70s^4 + 224s^3 + 147s^2 - 392s - 36)/144$
	$k = 5$		$(-5s^4 + 4s^3 + 21s^2 - 2s - 6)/24$	$(-7s^6 + 18s^5 + 85s^4 - 156s^3 - 264s^2 + 216s + 144)/144$
	$k = 6$		$(5s^4 - 15s^2 + 4)/120$	$(7s^6 - 12s^5 - 90s^4 + 80s^3 + 267s^2 - 36s - 72)/240$
	$k = 7$			$(-7s^6 + 6s^5 + 85s^4 - 20s^3 - 192s^2 + 8s + 48)/720$
	$k = 8$			$(7s^6 - 70s^4 + 147s^2 - 36)/5040$
$v = 2$	$k = 1$	$-s + 1$	$(-2s^3 + 6s^2 - 3s - 1)/12$	$(-3s^5 + 15s^4 - 10s^3 - 30s^2 + 24s + 4)/360$
	$k = 2$	$3s - 2$	$(10s^3 - 24s^2 - 3s + 16)/12$	$(21s^5 - 90s^4 - 20s^3 + 360s^2 - 213s - 54)/360$
	$k = 3$	$-3s + 1$	$(-10s^3 + 18s^2 + 15s - 15)/6$	$(-7s^5 + 25s^4 + 30s^3 - 130s^2 + 16s + 60)/40$
	$k = 4$	s	$(10s^3 - 12s^2 - 21s + 8)/6$	$(21s^5 - 60s^4 - 140s^3 + 336s^2 + 147s - 196)/72$
	$k = 5$		$(-10s^3 + 6s^2 + 21s - 1)/12$	$(-21s^5 + 45s^4 + 170s^3 - 234s^2 - 264s + 108)/72$
	$k = 6$		$(2s^3 - 3s)/12$	$(7s^5 - 10s^4 - 60s^3 + 40s^2 + 89s - 6)/40$
	$k = 7$			$(-21s^5 + 15s^4 + 170s^3 - 30s^2 - 192s + 4)/360$
	$k = 8$			$(3s^5 - 20s^3 + 21s)/360$

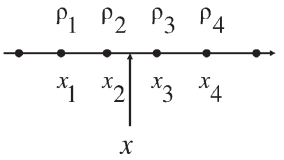


Fig. 9.1 One-dimensional example of the interpolation at x with $n = 4$ points, x_i and ρ_i are the abscissa and density values at grid point i.

x are at the center of the points arrangement. The interpolation is accomplished by a sequence of one-dimensional interpolations. For example, on a Cartesian mesh (Fig. 9.2) of tabulated values y_k (black circles) such as a parallelepiped, a two-dimensional square-mesh with approximated y_i (dark gray circles in Fig. 9.2)

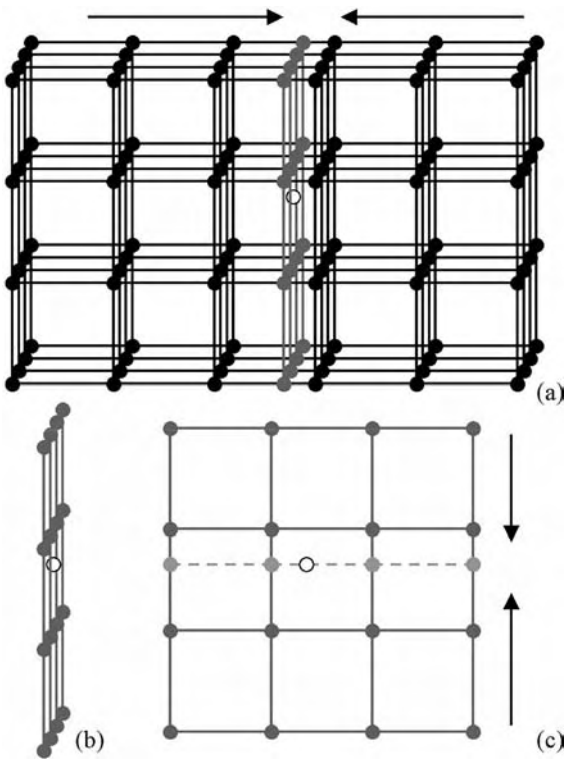


Fig. 9.2 Example of the interpolation method for a three-dimensional mesh of points. Black circles denote the points with tabulated values of the function. The open circle denotes the current point where the function must be interpolated. (a) A square of interpolated values at points denoted by gray circles is obtained applying the one-dimensional Lagrange interpolation to each line of six points along the black arrows. (b) and (c) show the square array after the first approximation. In (c) the light circles denote a new approximation using four-point linear interpolation along the black arrows. Finally, using a one-dimensional interpolation on the light gray line, the approximation of the function at the current value is obtained.

values on a plane that coincides with just one of the coordinates of the current point (open circle) is determined. Then, on the plane, a linear mesh (Fig. 9.2c) that coincides with an additional coordinate of the current point is approximated. Finally, the value of P is interpolated on this linear arrangement.

The final equation is:

$$P = \sum_{k_1=1}^{n_1} \sum_{k_2=1}^{n_2} \sum_{k_3=1}^{n_3} w_{k_1,n_1}(s_1) \cdot w_{k_2,n_2}(s_2) \cdot w_{k_3,n_3}(s_3) \cdot \gamma_{k_1 k_2 k_3} \tag{9}$$

Taking advantage of the crystal symmetry and the properties of the gradient paths of $\rho(\mathbf{r})$ associated with the CPs, we have implemented a very rapid, automated algorithm for systematic determination of the all CPs inside the unit cell of the crystal. First, it determines the bond CP among each pair of atoms (first and second neighbors) using the NR method. Because the gradient paths associated with the negative eigenvalues at the bond CPs originate mainly in the cage CPs, the algorithm then simply searches the origin of some of these paths (for each bond CP) applying the NR method at these points. Finally, searching along the lines connecting nearest neighbor cage CPs, all the ring CPs can be found. Symmetry is used to generate all the CPs having the same type of symmetry (same Wyckoff letter) and to avoid calculation of a CP that has already been determined.

9.3 An Application to Nanocatalyts – Exploring the Structure of the Hydrodesulfurization MoS_2 Catalysts

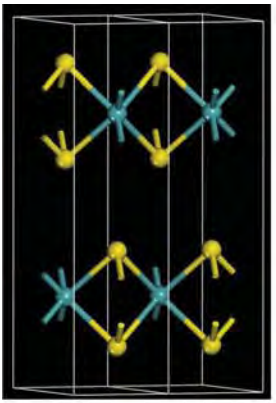
Transition metal sulfides (TMS) are a very important class of catalysts characterized by stability under harsh conditions in hydrodesulfurization (HDS), hydrodenitrogenation (HDN), and hydrogenation reactions [13–15]. In these processes the surfaces of the sulfides are reduced by sulfur elimination, by use of a large excess of hydrogen at temperatures ranging from 573 to 673 K creating coordinatively unsaturated sites (CUS) or vacancies around the metals. The CUS behave as electron-withdrawing sites whose properties may be regarded as a Lewis acid type center interacting with electron-donating organic substrates [16–18]. It is suggested the nature of these sites is intimately related to the metal–sulfur bond strength [16–19]. Basic studies support the idea that differences between catalytic activity is related to variations in the concentration of CUS (the Lewis acid sites), which in turn depend on the metal–sulfur bond strength [20–31]. Nickel (and cobalt)-promoted molybdenum sulfide catalysts have for many years been regarded as being among the most important catalysts used in refineries. Studies using X-ray absorption fine structure (EXAFS) have established that the active Mo atom is present as small MoS_2-like nanostructures [32, 33]. Adsorption and activity experiments [34, 35] have revealed that the active sites reside at the edges

of the MoS_2 structures, and high-resolution scanning tunneling microscopy (STM) and density-functional theory (DFT) studies [36] have recently shown that the MoS_2 nanoclusters adopt a hexagonal shape exposing two different types of edge, Mo edges covered with S monomers and fully sulfur-saturated S edges. Incorporation of nickel or cobalt into these edges significantly increases the activity of the catalyst [14, 37–39]. Under typical sulfidation conditions Ni is preferentially incorporated into the metal edge.

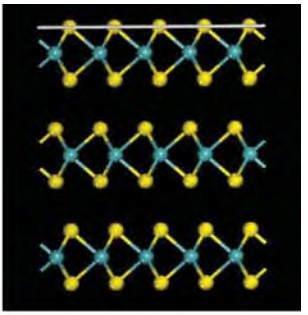
9.3.1 Catalyst Models

MoS_2 bulk is a layered-type crystal the lattice of which is described by the hexagonal space group P63/mmc with $a = b = 3.160\,Å$ and $c = 12.294\,Å$ [40]. Its crystal structure belongs to a family of polytypic structures with close-packed triangular double layers of S with Mo atoms arranged in the trigonal–prismatic holes of the S double layers (Fig. 9.3).

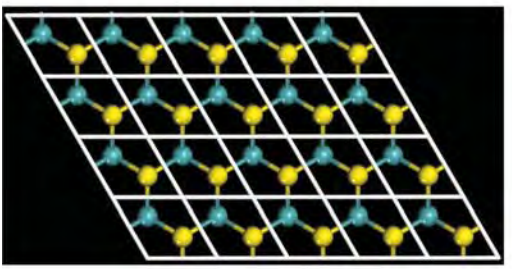
Mo atoms occupy the *2c* Wyckof positions with coordinates (1/3, 2/3, 1/4) and the S atoms the 4f position with coordinates (2/3, 1/3, 0.371) [41]. Each Mo atom is bonded to six S atoms in a trigonal–prismatic arrangement. The closest S–S distances are across the double layer and within the close-packed layers; the interlayer S–S distances are much larger and of the van der Waals type. The morphology of the catalysts [16] can be depicted as small MoS_2 particles (crystallites) dispersed at the surface of the support (usually SiO_2, graphite, etc). These particles have an average size of approximately $600 \pm 200\,Å^2$ and their reactivity depends on preferential exposed faces or planes. HDS catalysis is largely a surface process and, therefore, we must consider surface models. These models are usually obtained by cutting the three-dimensional bulk structure along a particular plane defined by using the appropriate Miller index [29, 39, 42]. For example, the so called basal plane of a MoS_2 crystallite [36] is produced by cleaving the crystal along the (001) plane (Figs. 9.3b and 9.3c). This plane is fully covered by sulfur atoms and is inactive for HDS reactions. Cleavage of the bulk structure parallel to (010) plane (Figs. 9.3d and 9.3e) produces the well known edge surface exposing coordinatively unsaturated molybdenum or sulfur atoms. Each of the exposed Mo atoms is coordinated to four sulfur atoms and each terminal sulfur atom is coordinated to two Mo atoms. Several studies have shown that the bare Mo edge terminating in a row of undercoordinated Mo atoms is very unfavorable; such edges will therefore have high affinity for S adsorption. The real morphology of the MoS_2 catalyst active sites has been deduced from experimental and theoretical studies. STM enables direct imaging of catalytically relevant surface structure on the atomic scale. By studying a realistic HDS model system consisting of a few-nanometer-wide gold-supported MoS_2 particles it has been shown that the morphology of the nanoparticles is sensitive to sulfiding and reaction conditions [36]; this means triangles are formed under heavy sulfiding conditions and truncated hexagons under more sulfo-reductive conditions resembling HDS conditions. These hexagonal clusters expose the basal plane and two different



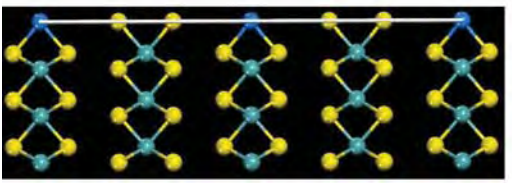
(a)



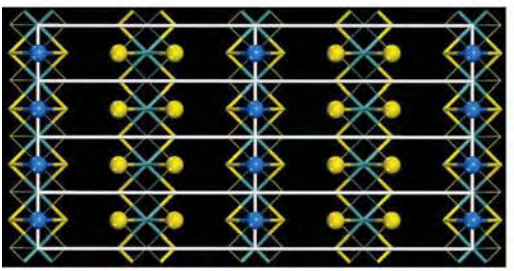
(b)



(c)



(d)



(e)

types of edge – Mo edges covered with S monomers and fully saturated S edges [36]. Different models have been used to study the edge structure of MoS_2 catalysts using DFT methods – cluster models including a finite number of atoms [43, 44], a single S–Mo–S periodic slab [45, 46], and a larger slab model containing two S–Mo–S sheets exposing Mo and S edges alternately [29, 38, 39, 47–54]. From these studies a clear picture of the MoS_2 edge structures has emerged. The most external Mo atoms of the Mo edge contain 50% sulfur coverage with each sulfur atom bridged to two neighboring molybdenum atoms. Significant reconstruction of the Mo edge occurs; the bridged S atoms are shifted by half a lattice constant relative to the bulk S lattice and move down to a bridging position in-plane with the Mo lattice. The S edges remain fully sulfided and a maximum Mo coordination to six sulfur atoms is achieved in one configuration where the edge is terminated by a row of sulfur atoms positioned in a bridge position close to those expected from bulk-terminated MoS_2. Both edge geometries lead to a coordination number of six for the outermost molybdenum atoms and two for the corresponding sulfur atoms. The local atomic structure of NiMoS catalysts has been resolved by means of DFT calculations coupled with simple thermodynamics determinations [29, 38]. Under typical sulfidation conditions nickel is preferentially incorporated into the metal edge in a square planar geometry with 0% sulfur. On a partially promoted metal edge, sulfur atoms bond to the outermost Mo atoms and the promoter atoms tend to be uncovered [38]. High-resolution electron microscopic studies of silica-supported Mo based catalysts have shown that the morphology of the catalyst can be depicted as small particles with an average size of 29 Å (mean diameter) and three slabs in width dispersed on the surface of the support [16]. Figure 9.4a shows a model of such a particle interacting with a dibenzothiophene molecule. The structure of the particle edge (the active site) is emphasized by means of a white square.

In this section we will study the larger periodic slab model (Fig. 9.4b) that exposes alternating layers of Mo and S edges denoting nanoparticles several layers wide. The unit cell ($9.480 \times 12.294 \times 36.000\,Å^3$) of this surface has three and six bridged S atoms above the Mo and S edges, respectively. This cell contains a periodic slab of several layers of atoms initially having the same structure of the surface built directly from the bulk. Vacuum layers thicker than 15 Å were used to ensure there were no interactions between adjacent slabs. The geometry of the models was optimized by using algorithms included in the Dmol^3 program [55, 56]. The two upper rows were allowed to relax while the atoms of the lower rows were kept fixed at their optimized bulk positions to simulate bulk constraints.

Fig. 9.3 (a) Ball-and-cylinder model of the MoS_2 bulk 1×2 cell illustrating the sheet arrangement of Mo and sulfur atoms. (b) Side view of three sheets and (c) Top view of a 4×5 cell of the (001) MoS_2 surface (the basal plane). (d) Side view and (e) top view of a 4×2 cell of the hypothetical (010) MoS_2 surface. White cylinders denote the unit cell whereas light blue and yellow spheres denote the Mo and S atoms, respectively. Spheres in (e) emphasize the outermost atoms of the surface.

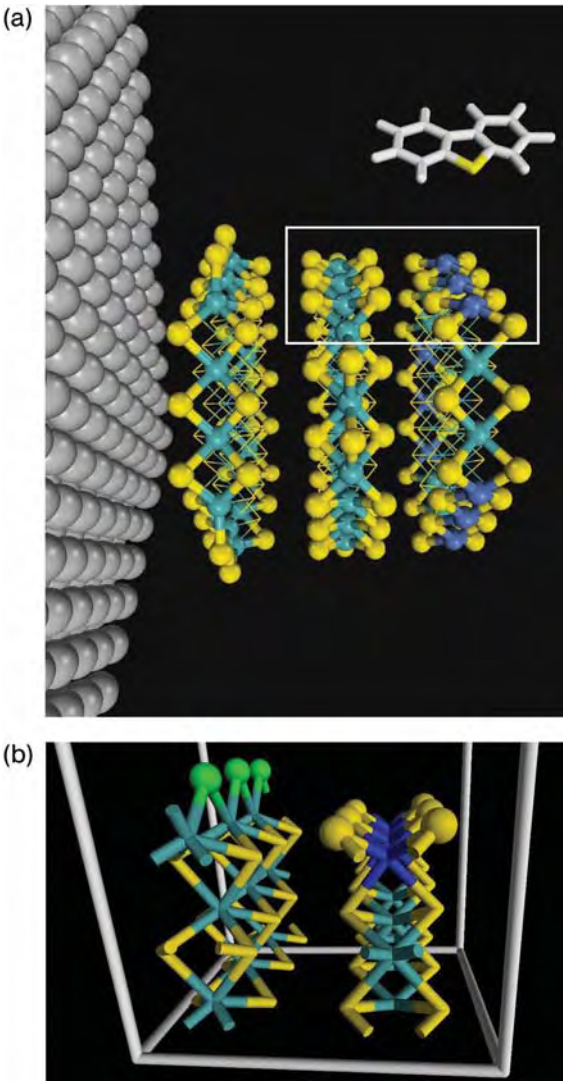


Fig. 9.4 (a) Ball-and-cylinder model showing a side view of a hypothetical NiMoS nanoparticle that simulates the particles of Ref. [16] supported on an Au (111) surface. Blue, dark blue, and yellow spheres denote the Mo, Ni, and S atoms, respectively, on the edges. A white rectangle emphasizes the structure of the active site that interacts with a dibenzothiophene (white and yellow cylinders) molecule. (b) Side view of the periodic model of the MoS_2 edges. Light blue and yellow cylinders denote the Mo and S atoms, respectively. Dark blue cylinders denote the Mo atoms on the S edges. Green and yellow spheres denote the outermost S atoms on the Mo and S edges, respectively.

9.3.2
The Full $\rho(\mathbf{r})$ Topology of the MoS_2 Bulk

In the MoS_2 unit cell, six sulfur atoms locally coordinate with one Mo atom to form a trigonal prismatic structure. Each Mo atom is surrounded by the six nearest sulfur atoms at a distance of 2.404 Å forming an MoS_2 sheet, and each S atom is surrounded by the three nearest Mo atoms on the sheet and by the three second nearest S atoms located on a neighboring MoS_2 sheet. All the CPs of that unit cell have been located and the data that characterize them are given in Table 9.2. In this table the corresponding Wyckoff letter in the International Tables of Crystallography [57] identifies the critical points within a unit cell. This identification is useful for determining the topology of the electron density of an extended system. Figure 9.5a illustrates the bond and cage CPS determined inside the MoS_2 sheets. The bond paths are shown as gray lines connecting the bound atoms. There are six nuclei in the primitive cell, two molybdenum (light blue spheres), located at the position labeled *c*, and four sulfur (yellow spheres), at position f. There are twelve Mo–S bond critical points (gray spheres) at position *k*, six four-membered ring CPs at *f*, and one trigonal prism-like cage (red spheres) at *b*.

Table 9.2 Topological properties (au) of $\rho(\mathbf{r})$ at the critical points for MoS_2 bulk space group: **P63/mmc** (D_{6h}^4 in Schoenflies notation).

Wyckoff letter	Site symmetry	Critical point	λ_1	λ_2	λ_3	ρ_b
***k* 12**	C_S	Mo–S *b*	−0.1073	−0.0809	0.3212	0.0908
***g* 6**	C_{2h}	S–S *b*	−0.0056	−0.0056	0.0388	0.0112
***h* 6**	C_{2v}	Four-membered 2Mo–2S *r*	−0.0347	0.0058	0.0492	0.0433
***b* 2**	D_{3h}	Three-membered 3 Mo *r*	−0.0149	0.0288	0.0292	0.0159
***k* 12**	C_S	Four-membered 1Mo–3S *r*	−0.0024	0.0062	0.0127	0.0070
***d* 2**	D_{3h}	Red *c* (Fig. 9.6b)	0.0253	0.0260	0.0664	0.0311
***f* 4**	C_{3v}	Green *c* (Fig. 9.6c)	0.0048	0.0062	0.0062	0.0057
***a* 2**	D_{3d}	Pink *c* (Fig. 9.6e)	0.0017	0.0022	0.0022	0.0045
***c* 2**	D_{3h}	Mo *n*				
***f* 4**	C_{3v}	S *n*				

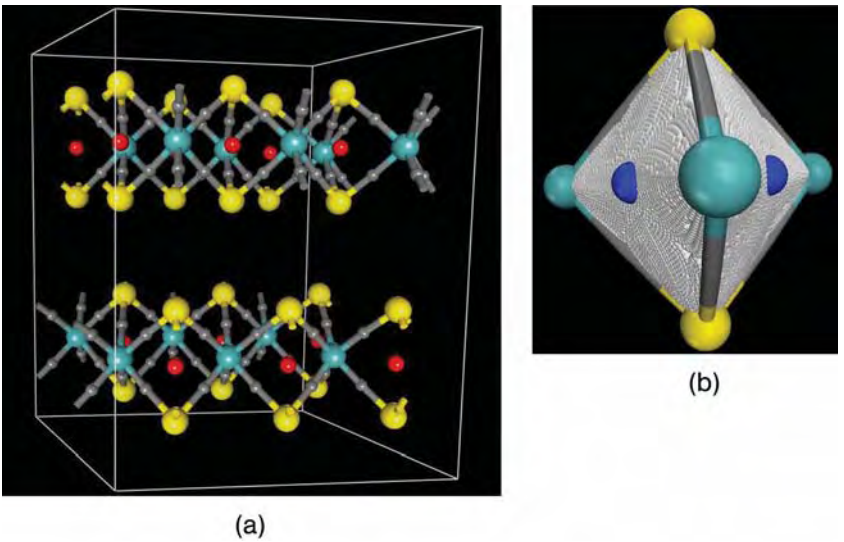
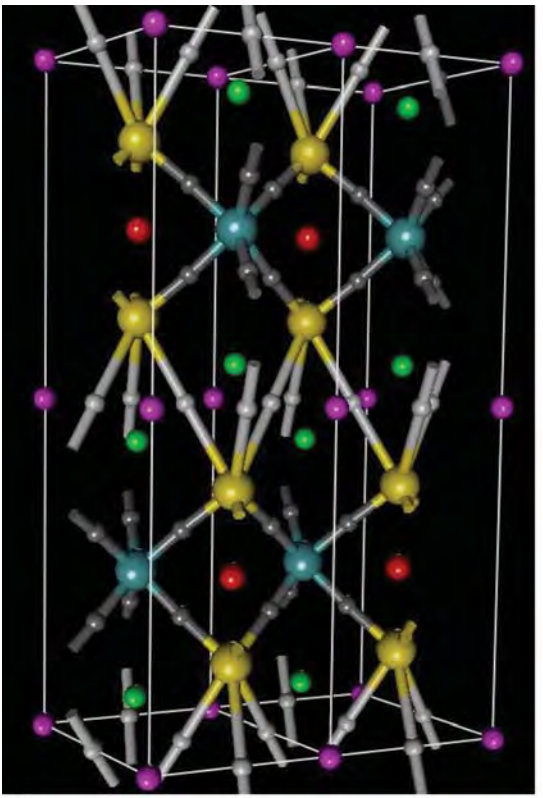


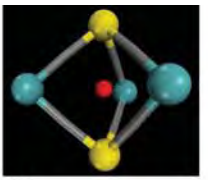
Fig. 9.5 (a) Bond (dark gray spheres) and cage (red spheres) $\rho(r)$ critical points located inside the bulk MoS_2 sheets. Large light blue and yellow balls denote the Mo and S atoms, respectively. (b) Representation of the ring surface (white faces) defined by the gradient paths that originate at the ring CPs (blue spheres) and end at the nuclei. Three of these ring surfaces define the faces of the red cages.

This cage is defined by just three nonplanar curved faces (Fig. 9.5b) with a ring CP at the center of each face. A set of four Mo–S bond paths delineates each of those ring structures. The set of gradient paths which originates at each of those ring CPs and terminate at the nucleus defines the so called ring surface (Fig. 9.5b). A set of three of those ring surfaces defines the cages located at *b*.

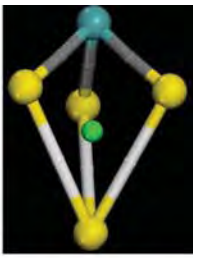
An additional set of six S–S bond CPs was also found at position g (white spheres in Fig. 9.6a) in the middle of each pair of neighboring MoS_2 sheets. These bond CPs are weaker interactions than the Mo–S bonds, as evidenced by the values of ρ ($\rho_{Mo-S} = 0.0908$ au and $\rho_{S-S} = 0.0112$ au) at $\mathbf{r_c}$ and are characteristic of van der Waals interactions [1, 10, 58–60]. Similar results have been published for many hydrogen-bonded systems [59, 60] and van der Waals molecules [58]. The presence of these S–S bond CPs creates additional sets of ring and cage points between the MoS_2 sheets (Figs. 9.6a and 9.6c). Four distorted trigonal prism cages (green spheres) at *f* positions, two fourteen-faced cages at *a* positions (pink spheres), twelve four-membered rings at *k* and two six-membered ring CPs at *d* have also been determined. Each cage at *f* is formed by three curved faces (Fig. 9.6c) defined by two Mo–S and two S–S bond paths with a ring CP at the center of each face. The cages located at *a* positions are defined by fourteen curved faces (Figs. 9.6d and 9.6e): six four-membered faces shared



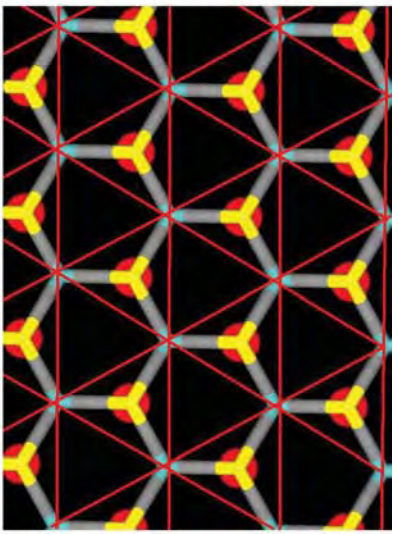
(a)



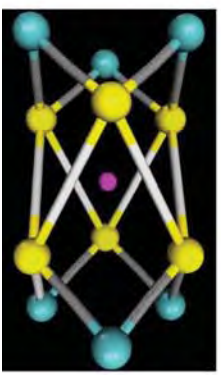
(b)



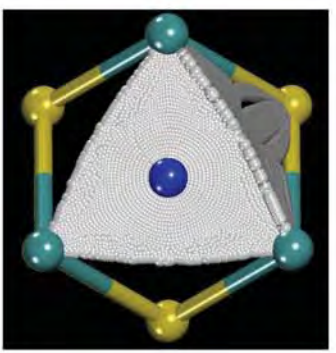
(c)



(d)



(e)



(f)

with the cages at *f* (green spheres), two three-membered faces defined by a ring surface bonding three Mo atoms at the top and at the bottom of the cage, and six three-membered faces shared (just the half of the face) with the cages at *b* (Fig. 9.6f).

The entire set of critical points defines specific bonding polyhedra, the packing of which gives rise to the MoS_2 bulk structure. The topology of these polyhedra is determined solely by the number of corners, edges, and faces. Convenient notation for these polyhedra [5] is {corner number, edge number, face number}. The network of *bond paths* (Figs. 9.6a and 9.6d) describing the atomic connectivity shows that the crystal graph of MoS_2 results from packing of three types of polyhedron – two {5, 6, 3} such as that shown in Figs. 9.6b and 9.6c and one {12, 18, 8} (Fig. 9.6e), with each face of the distorted trigonal prisms located at *f* shared with a {12, 18, 8}. The volume spanned by the paths ending at a given nucleus defines the basin of the atom. A 3D view of the IAS and the atom basin are given in Fig. 9.7. There are eleven cage critical points defining an Mo atom and a set of eight cages defines the S atoms. The paths that originate at four cages and terminate at a *b* CP define each inter-atomic surface. There is a set of six Mo–S surfaces bounding the Mo atomic basin and a set of three Mo–S and three S–S surfaces bounding the S atoms. Consequently, the Mo and S atoms have D_{3h} and C_{3v} symmetry, respectively. Along the Mo–S bond paths, both atoms have curved faces corresponding to Mo–S bonds – concave in the basin of the molybdenum and convex in that of the sulfur. This result agrees with a transfer of electronic density from the Mo atom to the sulfur atom [1, 31]. These basins display the full local point-group symmetry at the nuclear sites and fill the space without overlapping each other.

Fig. 9.6 (a) Ball-and-cylinder model illustrating the bond and cages $\rho(r)$ critical points of the MoS_2 bulk. Large light blue and yellow balls denote the Mo and S atoms, respectively. Dark gray and red spheres denote the Mo–S bond and cage CPs, respectively, inside the sheets. White, gray, and pink spheres denote the S–S bond and cage CPs locate between the sheets. (b) Five nuclei (three Mo and two S), six Mo–S bond paths, and three four-membered curved faces define each {5, 6, 3} red cage. (c) Green cages located at *f* Wyckoff positions define a layer of {5, 6, 3} polyhedra just below the red cages. (d) Top view (just the 001 Miller plane) of the cell shown in (a) illustrating the space (red triangle with a red sphere at the center) occupied by the {5, 6, 3} polyhedra defined by the red cages. The empty triangles (without red spheres at the center) are filled for {12, 18, 14} polyhedra defined by the pink cages. (e) Twelve nuclei (six Mo and six S), twelve Mo–S and six S–S bond paths, and fourteen faces define each {12, 18, 14} pink cage. (f) Representation of the three-membered ring surface (white faces) defined by the gradient paths that originate at the ring CPs (dark blue spheres) and end at the Mo nuclei. Three lateral three-membered ring surfaces (gray cylinders) complete the top of the cage. Right: half of the ring surfaces shared with the red cages define these lateral three-membered faces.

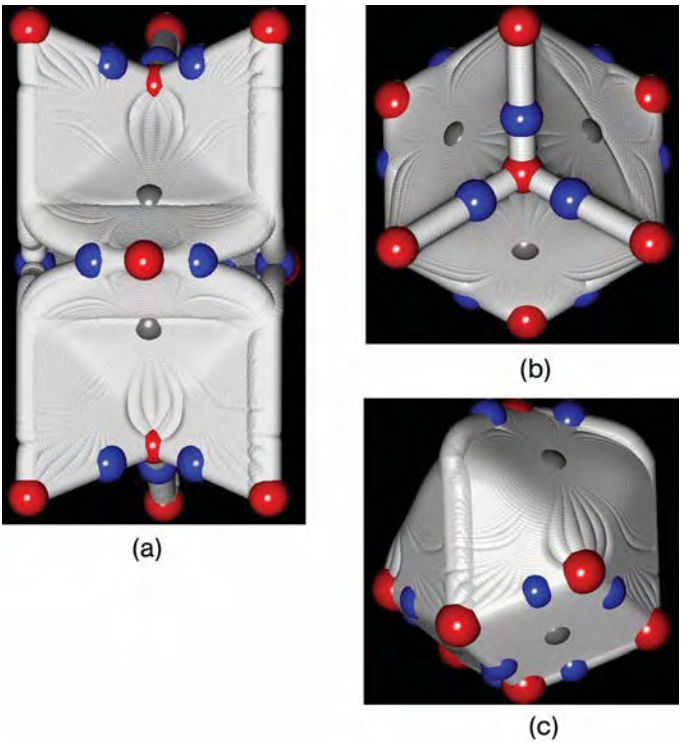


Fig. 9.7 (a) Side and (b) top view of the region of space (the basin) bounded by six concave interatomic surface that defines an Mo atom. Each interatomic surface is defined by the paths that originate at four cages CP (red spheres) and terminate at Mo–S bond CPs (gray spheres). (c) Side view of the S atom basin. Three convex (at the top) and three planar (at the bottom) interatomic surfaces are defined by the paths that terminate at three Mo–S and three S–S bond CPs, respectively. Dark blue spheres denote ring CPs that bound each pair of cage CPs.

9.3.3
The $\rho(\mathbf{r})$ Topology of the MoS_2 Edges

The $\rho(\mathbf{r})$ topology of the MoS_2 edge model (Fig. 9.8) has shown the presence of almost the same type of CP as that of the bulk case – Mo–S (gray spheres) and S–S bond critical points (white cylinders), four-membered ring CPs, trigonal prism-like cages (small red spheres), distorted trigonal prism cages (green spheres), and fourteen-faced cages (pink spheres). A new kind of cage (big red spheres) capped by the outermost sulfur atoms on the metal edge (big yellow spheres) is the main difference from the bulk CPs. These cages are built with four curved faces and describe {6, 8, 4} polyhedra (Fig. 9.8b). Thus, the graph of the exposed structure for the MoS_2 surface results from the packing of these {6, 8, 4} and {5, 6, 3} polyhedra similar to the bulk one with a green cage at the

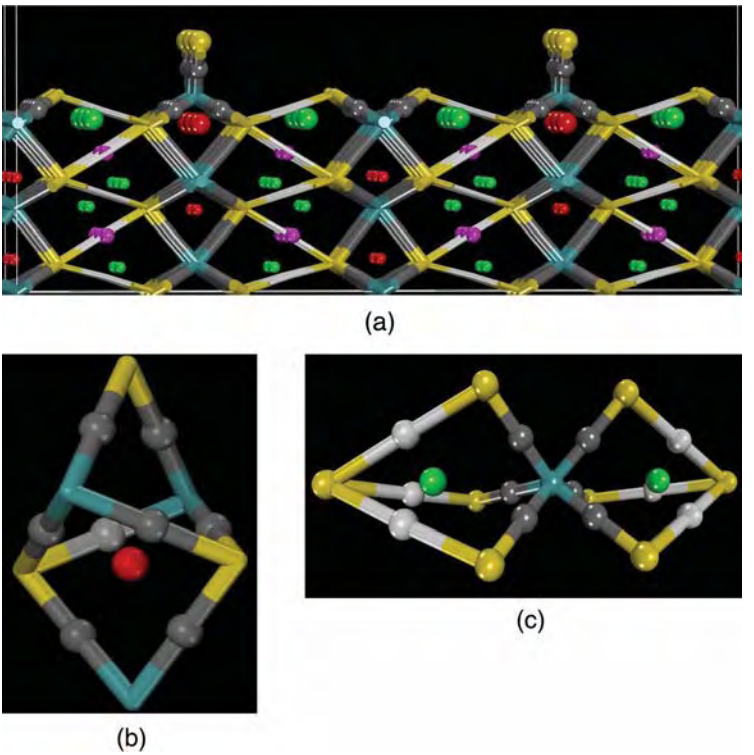


Fig. 9.8 (a) Side view of a ball-and-cylinder model illustrating the bond and cages $\rho(r)$ critical points of the MoS_2 edge. Large light blue and yellow balls denote the Mo and S atoms, respectively. Dark gray spheres and white cylinders denote the Mo–S and S–S bond CPs, respectively. Red, green, and pink spheres denote the cage CPs in the same sense of Fig. 9.6. Big spheres emphasize the outermost cage CPs. (b) {6, 8, 4} polyhedron defined for the outermost S atoms on the metal edge. (c) Most external {5, 6, 3} polyhedra bonded to the exposed Mo atoms of the sulfur edge.

middle (Fig. 9.8c). Each {6, 8, 4} polyhedron is defined by eight Mo–S bond paths with an average ρ_b value (Table 9.3) of 0.0926 au, and contains a cage CP with a ρ_c value of 0.0297 au, and the {5, 6, 3} polyhedra are defined by three Mo–S bond paths with an average ρ_b value of 0.0916 au and three S–S bond with ρ_b value of 0.0087 au and contain a cage with ρ_c value of 0.0054 au. The former polyhedra whose nuclei are bonded by a much bigger electron density should therefore be more stable than the {5, 6, 3} polyhedra.

The basin for the outermost Mo atoms is bordered by six Mo–S inter-atomic surfaces. A set of seven cages (the paths that originate at the cages and terminate at the Mo nuclei outline its associated basin) defines these Mo atoms. In a similar way to the bulk example, the internal IAS are generated by the paths that originate at four cages while the outermost IAS are defined by two cages only and by the paths from the infinite (the open space) and terminate at the bond CPs.

Table 9.3 Topological properties (au) of $\rho(\mathbf{r})$ at the outermost critical points for MoS_2 edges.

Edge	Critical point	λ_1	λ_2	λ_3	ρ_b
Sulfur	Mo–S *b*	−0.1011	−0.0943	0.3054	0.0916
Sulfur	Four-membered 2Mo–2S *r*	−0.0068	0.0216	0.0496	0.0323
	Three-membered *r*				
Sulfur	Four-membered 1Mo–3S *r*	−0.0023	0.0072	0.0127	0.0061
Sulfur	Red *c* (Fig. 9.8)	0.0066	0.0122	0.0733	0.0311
Both	Green *c* (Fig. 9.8)	0.0015	0.0051	0.0087	0.0054
Both	Pink *c* (Fig. 9.8)	0.0012	0.0026	0.0026	0.0029
Both	S–S *b*	−0.0047	−0.0043	0.0307	0.0087
Mo	Mo–S *b* new	−0.1010	−0.0962	0.3215	0.0926
Mo	*c* new (Fig. 9.8)	0.0050	0.0117	0.0690	0.0297

The basin for the outermost S atoms on the Mo edge and on the sulfur edge is bordered by two and four Mo–S IAS, respectively. Figure 9.9 shows a top view of several of those S basins.

It is apparent the Mo atoms are totally covered by the sulfur atoms, which obstruct access of incoming molecules, for example dibenzothiophene that contains polluting sulfur atoms [47, 50, 52, 54]. This result agrees with the known inactivity of this kind of site, similar the basal plane that only exposes sulfur atoms. We must therefore create vacancies or CUS on this surface by a reduction process to generate available Lewis acidic sites. Usually, H_2 reacts with surface sulfur atoms to create a vacancy and produce H_2S [50]. The energy to create a sulfur vacancy can therefore be calculated by using the equilibrium: Surface-S + H_2 ⇔ Surface-□ + H_2S (where '□' denotes a vacancy). Thus the creation energy, CE, of a sulfur vacancy is given by the expression:

$$\mathrm{CE} = \mathrm{E(Surface\text{-}€)} + \mathrm{E}(H_2S) - \mathrm{E(Surface\text{-}S)} - \mathrm{E}(H_2) \tag{10}$$

The energy for each sulfur removal has been reported and the stability of each surface has been deduced using a DFT methodology based in pseudo potentials and plane wave basis sets [51]. The CE and the Lewis acidity strength of sulfur vacancies or coordinative unsaturated sites on the MoS_2 edges were recently studied using density-functional theory for periodic systems and an electrostatic potential based method [61]. On the sulfur edge, the gradual removal of the six

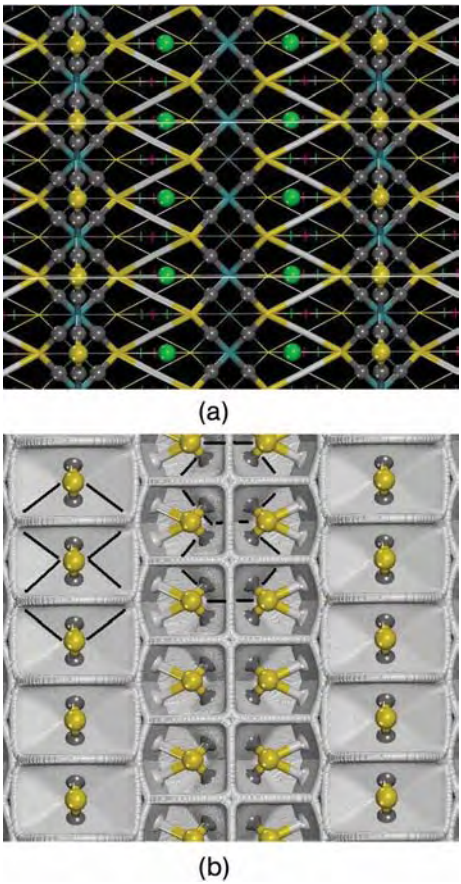


Fig. 9.9 (a) Top view of a ball-and-cylinder model illustrating the bond and cages $\rho(r)$ critical points of the MoS_2 edge. (b) Top view of the MoS_2 edge surface showing the IAS defining the exposed basin of the outermost S atoms. Black lines emphasize the basin of some Mo atoms, showing that these atoms are encapsulated by the external S atoms.

S atoms of the cell (Fig. 9.4b) was studied and for each case geometrical optimization was conducted for the most stable configuration reported. For the Mo edge, gradual removal of the three S atoms (Fig. 9.4b) ending with complete exposure of the underlying Mo atoms was also studied. The Lewis acidity strength of these specific sites can be explored by use of the electrostatic potential, $V(\mathbf{r})$, which enables us to determine directly where the electron-rich sites in a molecule or crystal are located [61–72].

$V(\mathbf{r})$ at a point $\mathbf{r}$ generated by a molecule or crystal is given by:

$$V(\mathbf{r}) = V_N(\mathbf{r}) + V_E(\mathbf{r}) \tag{11}$$

where the two terms $V_N(r)$ and $V_E(r)$ represent the bare nuclear and electronic contributions, respectively, to the total electrostatic potential. The sign of $V(r)$ at a given point indicates either the nuclear (positive) or the electronic (negative) effects are dominant. The electrostatic potential at r generated by the total charge distribution, ρ^{tot}, of a periodic system is given by:

$$V(\mathbf{r}) = \sum_n \int \rho^{tot}(\mathbf{r}' - \mathbf{R}_n)|\mathbf{r} - \mathbf{r}'|^{-1}\, d\mathbf{r}' \tag{12}$$

The summation extends to all direct lattice vectors, the prime on the integral sign indicating that an infinitesimal region about $\mathbf{r} = \mathbf{r}'$ is excluded from the domain of integration to avoid divergent nuclear self-interaction terms that would otherwise arise in the electrostatic energy per cell. ρ^{tot} may be decomposed into electronic and nuclear components, as follows:

$$\rho^{nuc}(\mathbf{r}) = \sum_a q_a \delta(\mathbf{r}_a, \mathbf{r}) \tag{13}$$

where the summation extends to all the reference cell nuclei, with atomic numbers and position vectors denoted q_a and $\mathbf{r}_a$, respectively.

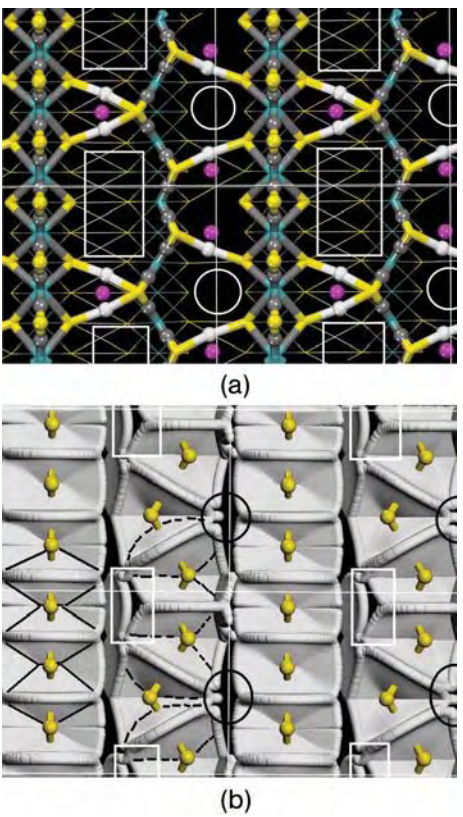
$$\rho^{el}(\mathbf{r}) = -\sum_{ij} \sum_{\mu\nu} \mathbf{P}_{\mu\mathbf{R}_i\nu\mathbf{R}_j} \chi_\mu(\mathbf{r} - \mathbf{R}_i)\chi_\nu^*(\mathbf{r} - \mathbf{R}_i) \tag{14}$$

where P is the density matrix and $\chi_\mu(\mathbf{r} - \mathbf{R}_i)$ is the μ_{th} reference cell basis function translated by the direct lattice vector $\mathbf{R}_i$. The summations over i and j extend to all direct lattice vectors, while those over μ and ν include all the basis functions of the reference cell. Substitution of Eq. (13) and Eq. (14) into Eq. (12) gives the nuclear and electronic $V(\mathbf{r})$ contributions.

For the region nearest to the nucleus V_N dominates and the topology of $V(r)$ is similar to that of the electron density [73, 74], $\rho(r)$, i.e. positive maxima at the nuclear site and a positive saddle point between every pair of bonded atoms. The existence of maxima is, nevertheless, ruled out by an established result that, except for the nuclear position, there cannot be any strict local maxima in the $V(r)$ map [65, 66]. For the region where V_E dominates, $V(r)$ is negative and the $V(r)$ topography can be more complex. It is, however, well known that lone pairs of electrons and double π bonds (C=C, C=N, etc) are usually characterized as negative minima [64, 66]. In summary, the region nearest the nucleus is always positive whereas the region where the potential is negative contains the minima that characterize the atom lone pairs. The minima of the negative region denote the zones to which an approaching electrophile may be attracted. In contrast, the positive regions do not have maxima that might indicate sites for nucleophilic attack. Politzer and Sjoberg have, nevertheless, shown that by computing $V(r)$ on the

Table 9.4 Creation energy, CE (Eq. 10), and maximum positive value of $V(r)$ for the sites studied on the Mo and S edges of molybdenum sulfide [61].

Site	Edge	Number of sulfur atoms removed	CE (kJ mol^{-1})	V(r) maximum (kJ mol^{-1})
2	S	2	34.581	15.491
3	S	3	72.579	51.982
4	S	4	260.579	54.085
5	S	5	515.409	131.695
6	Mo	1	153.345	44.371
7	Mo	2	271.810	142.040
8	Mo	3	458.538	142.877



0.002-electron $Å^{-3}$ contour isosurface [75] of the molecular electronic density $\rho(\mathbf{r})$, the susceptibility of molecules to nucleophilic attack can be quantified. They demonstrated that the relative magnitudes of the positive electrostatic potential in different regions on this surface do reveal the sites most susceptible to nucleophilic attack. This contour isosurface for a group of diatomic molecules and for methane encompasses at least 95% of the electronic charge and yields physically reasonable molecular dimensions [75]. A similar approach, mapping the $V(\mathrm{r})$ values on this isosurface on to colors to identify the host sites in which nucleophiles (most positive zone) and electrophiles (most negative zone) should bind, was used to explore the Lewis acid sites on the MoS_2 edges. In addition, to quantify the susceptibility of the active sites, the minimum and maximum $V(\mathrm{r})$ values at the determined host zones were also determined using the Newton–Raphson technique described in Section 9.3. The reported CEs and $V(\mathrm{r})$ maximum values [61] that characterize the Lewis acid strength of the studied sites are listed in Table 9.4. These values suggest that the energetically more favorable sites are located on the sulfur edges; their Lewis acidity strength is, however, much smaller than the site acidity at the molybdenum edges. As a compromise of Lewis acidity strength and energy site 3 with 50% sulfur coverage is suggested as the most hydrodesulfurization-active site. Sites 5, 7, and 8 are very acid sites but at a cost of a huge CE, as is apparent from Table 9.4. These CE results confirm the previously suggested greater stability of the {6, 8, 4} polyhedra on the Mo edges than the {5, 6, 3} polyhedra on the S edges. Because removal of each S atom of the S edge destroys two {5, 6, 3} polyhedra, destruction of six these last polyhedra to create site 3 costs half the energy necessary to destroy just one {6, 8, 4} polyhedron to create site 6 (Table 9.4).

The $\rho(\mathrm{r})$ topology for site 3 is shown in Fig. 9.10a. This site is obtained by removing three S atoms from the cell, just on the S edge; the remaining three S atoms again occupy bridging positions between the Mo sites. In contrast with the Mo edge, however, the S atoms are tilted to conform as far as possible to the bonding Mo–S network present in the solid. The Mo–S distances at the S edge vary between 2.29 and 2.32 and are shorter than the Mo–S distances (2.40) of the surface without vacancies. The S atoms with the shortest Mo–S distances (two by cell) have only one S–S bond (Fig. 9.10a) completing tricoordinated S atoms whereas the other S atom of the cell preserves the two S–S bonds located

Fig. 9.10 (a) Top view of a ball-and-cylinder model illustrating the bond and cage $\rho(\mathrm{r})$ critical points of site 3. Large light blue and yellow balls denote the Mo and S atoms, respectively. Dark gray spheres and white cylinders denote the Mo–S and S–S bond CPs, respectively. Pink spheres denote cage CPs in the same sense as in Fig. 9.8. Note that holes (emphasized by white rectangle and circles) have appeared in the bond path network. (b) Top view of site 3 showing the IAS defining the exposed basin of the outermost S atoms. Holes in the interatomic surfaces (emphasized by white rectangle and black circles) suggest the site where the Mo atoms are most accessible to the incoming molecules. Dotted black lines show the basin of the Mo atoms on the sulfur edges just encapsulated by the external S atoms.

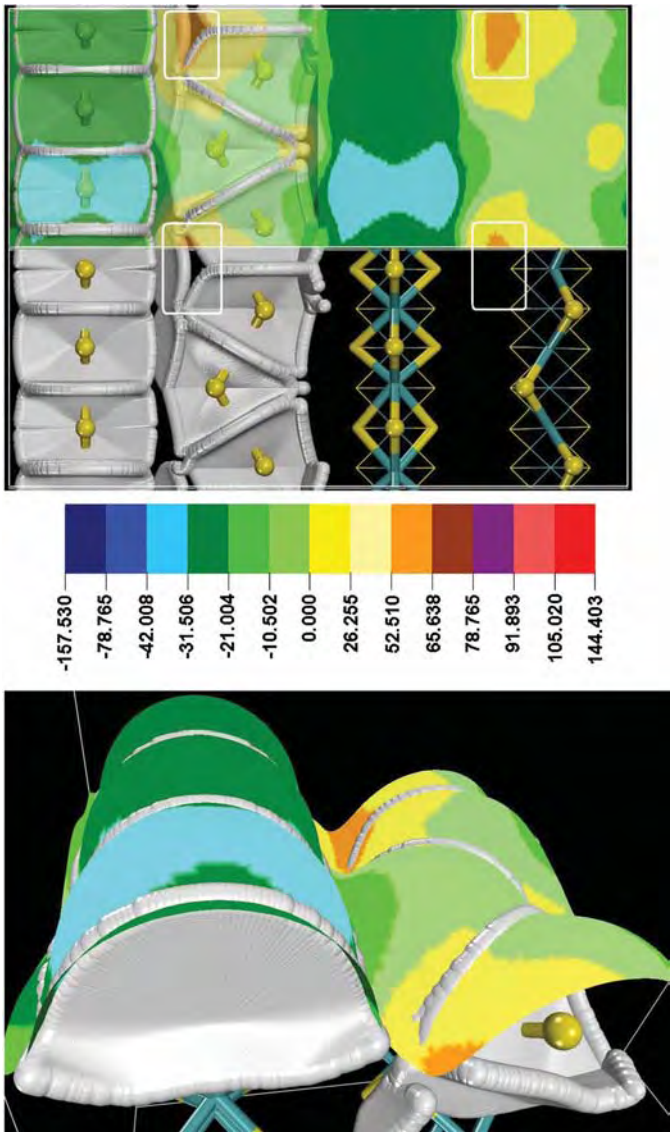


Fig. 9.11 (a) Top view of the 0.002 e Å^{-3} contour of $V(r)$ mapping of the $V(r)$ field values on to colors placed over the IAS defining the basin of the outermost sulfur atoms. The color map tab at the bottom is: dark blue (−157.530 to −78.765 kJ mol^{-1}), sky blue (−78.765 to 42.008 kJ mol^{-1}), light blue (−42.008 to −31.506 kJ mol^{-1}), dark green (−31.506 to −21.004 kJ mol^{-1}), green (−21.004 to −10.502 kJ mol^{-1}), light green (−10.502 to 0.000 kJ mol^{-1}), yellow (0.000 to 26.255 kJ mol^{-1}), yellow cream (26.255 to 52.510 kJ mol^{-1}), orange (52.510 to 65.638 kJ mol^{-1}), brown (65.638 to 78.765 kJ mol^{-1}), purple (78.765 to 91.893 kJ mol^{-1}), light red (91.893 to 105.020 kJ mol^{-1}) and red (105.020 to 144.403 kJ mol^{-1}). White squares indicate the most positive zones on the $V(r)$ contour. (b) Side view of the basin of the outermost sulfur atoms. The 0.002 au contour of $V(r)$ is used to define the border of these open atoms. $V(r)$ mapped on this border enables us to locate the stronger Lewis acid sites.

in the inter-edge space, forming four-coordinated atoms. Holes (denoted as a white rectangle in Fig. 9.10a) can be observed in the bond path network, between each pair of consecutive four-coordinated S atoms.

The basins for the outermost S atoms of site 3 are shown in Fig. 9.10b. It is apparent that compared with the surface without vacancies (Fig. 9.9b), the basins on the S edge have suffered a drastic change in their shape. These atoms spread to occupy the space liberated in the vacancies and adopt a triangular form like an arrow head. The base of each triangle is defined by the S–S IAS and the opposite angle (head point) penetrates inside the vacancy, occupying the space liberated by the removed S atoms, ending at the middle of the network bond path hole. The region around these head points (denoted by rectangles and circles in Fig. 9.10b) should be the sites with smaller sulfur electronic density on the surface, enabling access to the Mo atoms (the Lewis acid sites). This can be easily checked by means of the electrostatic potential. This let us visualize directly the stronger Lewis acid sites (most positive zones) by surface mapping the $V(\mathrm{r})$ values on to colors on the 0.002 electron Å^{-3} isosurface of $\rho(\mathrm{r})$. Figure 9.11a shows the superimposition of this color map on the exposed S atom basins of the MoS_2 edges model. Starting from the most negative $V(\mathrm{r})$ values (caption of Fig. 9.11a), three kinds of blue and three kinds of green denote the most negative values and two kinds of yellow and one orange denote the positive $V(\mathrm{r})$ values. The orange region corresponds to the most positive site and, as is apparent from Fig. 9.11, it perfectly matches with the hole where two sulfur basin head points almost converge.

Exposed atoms, atoms in an isolated molecule, have substantial open parts that extend to infinity. These atoms are open or unbounded at the exterior of the surface and a practical definition [9] is to cap the atom with an isosurface of the electron density with small $\rho(\mathbf{r})$ value representing the van der Waals envelope of the system. The 0.002 electron Å^{-3} contour of $\rho(\mathrm{r})$ shown in Fig. 9.11 just defines the border of the outermost sulfur atoms of the (010) MoS_2 surface. Figure 9.11b shows a side view of the atomic shape of these sulfur atoms. It is worthy of note that the stronger Lewis acid sites are located between the sulfur atoms at the bottom of the $\rho(\mathrm{r})$ 0.002 contour valley. This is the region where the Mo atoms are most accessible to incoming molecules.

In summary, there is a profound relationship between the vacancy creation energy, i.e. the energy necessary to remove exposed sulfur atoms from the MoS_2 surface and the nature (number of bond paths and $\rho(\mathbf{r})$ value at the bond and cage CPs) of the exposed polyhedra defining the surface graph. The structure and shape of the basins of the exposed sulfur atoms also enable us to visualize and locate the Lewis acid sites on this surface.

Acknowledgments

This work was supported by the grant G-2000001512 from the CONICIT (Consejo Nacional de Investigaciones Científicas y tecnológicas) of Venezuela.

References

1 Bader, R. F. W. *Atoms in Molecules – A Quantum Theory* (Clarendon Press: Oxford, U.K., **1990**).
2 Bader, R. F. W. *J. Phys. Chem.* **1998**, 102, 7314–7323.
3 Bader, R. F. W.; Popelier, P. L. A.; Keith, T. A. *Angew. Chem., Intl. Ed., Eng.* **1994**, 33, 620–631.
4 Zou, P. F.; Bader, R. F. W. *Acta Crystallogr.* **1994**, *A*50, 714–725.
5 Eberhart, M. E. *Can. J. Chem.* **1996**, 74, 1229–1235.
6 Eberhart, M. E. *Philos. Mag. A.* **1996**, 73, 47–60.
7 Aray, Y.; Rodriguez, J.; Vega, D. *J. Phys. Chem. B* **2000**, 104, 4608–4612.
8 Martín Pendás, A.; Costales, A.; Luaña, V. *Phys. Rev. B* **1997**, 55, 4275–4284.
9 Popelier, P. L. A. *Atoms in Molecules – An Introduction* (Prentice Hall, Harlow, UK, **2000**).
10 Gillespie, R. J.; Popelier, P. L. A. *Chemical Bonding and Molecular Geometry: From Lewis to Electron Densities* (Oxford University Press, New York, Oxford, **2001**).
11 *Numerical Recipes in Fortran 77: The Art of Scientific Computing.* (Cambridge University Press, **1992**). p. 355.
12 *Numerical Recipes in Fortran 77: The Art of Scientific Computing.* (Cambridge University Press, **1992**). p. 708.
13 Weisser, O.; Landa, S. *Sulfide Catalysts: Their Properties and Applications*, Pergamon, Oxford, UK., **1973**.
14 Topsoe, H.; Clausen, B. S.; Massoth, F. E. *Hydrotreating Catalysis Science and Technology, Vol. 11*, Springer Verlag, Berlin, **1996**.
15 Rodriguez, J. A.; Dvorak, J.; Capitano, A. T.; Gabelnick, A. M.; Gland, J. L. *Surf. Sci.* **1999**, *429*, L462–L468.
16 Berhault, G.; Lacroix, M.; Breysse, M.; Maugé, F.; Lavalley, J.-L.; Nie, H.; Qu, L. *J. Catal.* **1998**, *178*, 555–565.
17 Breysse, M.; Berhault, G.; Kasztelan, S.; Lacroix, M.; Maugé, F.; Perot, G. *Catal. Today.* **2001**, *66*, 13–20.
18 Pis Diez, R.; Jubert, A. H. *J. Mol. Catal.* **1993**, *83*, 219–325.
19 Nørskov, J. K.; Clausen, B. S.; Topsøe, H. *Catal. Lett.* **1992**, *13*, 1–8.
20 Toulhoat, H.; Kresse, G. *Am. Chem. Soc. Div. Petrol. Chem. Prepr.* **1997**, *42*, 114.
21 Pecoraro, T. A.; Chianelli, R. R. *J. Catal.* **1981**, *67*, 430–445.
22 Chianelli, R. R.; Daage, M.; Ledoux, M. J. *Adv. Catal.* **1994**, *40*, 117–232.
23 Harris, S.; Chianelli, R. R. *J. Catal.* **1984**, *86*, 400–412.
24 Vissers, J. P. R.; Groot, C. K.; Vanoers, E. M.; de Beer, V. H. J. *Bull. Soc. Chim. Belg.* **1984**, *93*, 813–821.
25 Ledoux, M. J.; Michaux, O.; Agostini, G.; Panisso, P. *J. Catal.* **1986**, *102*, 275–288.
26 Burdett, J. K.; Chung, J. T. *Surf. Sci. Lett.* **1990**, *236*, L353–L357.
27 Smith, T. S.; Johnson, K. H. *Catal. Lett.* **1994**, *28*, 361.
28 Orita, H.; Uchida, K.; Itoh, N. *Appl. Catal. A – Gen.* **2004**, *258*, 115–120.
29 Raybaud, P.; Hafner, J.; Kresse, G.; Kasztelan, S.; Toulhoat, H. *J. Catal.* **2000**, *190*, 128–143.
30 Neurock, M.; van Santen, R. A. *J. Am. Chem. Soc.* **1994**, *116*, 4427–4439.
31 Aray, Y.; Rodriguez, J.; Vega, D.; Rodriguez-Arias, E. N. *Angew. Chem. Int. Ed.* **2000**, *39*, 3810–3813.
32 Clausen, B. S.; Lengeler, B.; Candia, R.; Als-Nielsen, J.; Topsoe, H. *Bull. Soc. Chim. Belg.* **1981**, *90*, 1249–1259.
33 Parham, T. G.; Merrill, R. P. *J. Catal.* **1984**, *85*, 295–310.
34 Tauster, S. J.; Pecoraro, T. A.; Chianelli, R. R. *J. Catal.* **1980**, *63*, 515–519.
35 Topsoe, H.; Candia, R.; Topsoe, N. Y.; Clausen, B. S. *Bull. Soc. Chim. Belg.* **1984**, *93*, 783–806.
36 Lauritsen, J. V.; Bollinger, M. V.; Lægsgaard, E.; Jacobsen, K. W.;

Nørskov, J. K.; Clausen, B. S.; Topsøe, H.; Besenbacher, F. *J. Catal.* **2004**, *221*, 510–522.

37 Prins, R.; de Beer, V. H. J.; Somorjai, G. A. *Catal. Rev. – Sci. Eng.* **1989**, 31, 1–41.

38 Sun, S. M.; Adjaye, J.; Nelson, A. E. *App. Catal. A: General* **2004**, *263*, 131–143.

39 Sun, M.; Adjaye, J.; Nelson, A. *J. Catal.* **2004**, *226*, 32–40.

40 Bronsema, K. D.; De Boer, J. L.; Jellinek, F.; *Z. Anorg. Allg. Chem.* **1986**, *540*, 15–17.

41 Raybaud, P.; Kresse, G.; Toulhoat, H. *J. Phys. Condens. Matter.* **1997**, *9*, 11085–11106.

42 Aray, Y.; Rodríguez, J.; Vega, D.; Coll, S.; Rodríguez-Arias, E.; Rosillo, F. *J. Phys. Chem. B.*, **2002**, *106*, 13242–13249.

43 Byskov, L. S.; Norskov, J. K.; Clausen, B. S.; Topsoe, H. *Catal. Lett.* **2000**, *64*, 95–99.

44 Schweiger, H.; Raybaud, P.; Toulhoat, H. *J. Catal.* **2002**, *212*, 33–38.

45 Byskov, L. S.; Norskov, J. K; Clausen, B. S.; Topsoe, H. *Catal. Lett.* **1997**, *47*, 177–182.

46 Byskov, L. S.; Norskov, J. K.; Clausen, B. S.; Topsoe, H. *J. Catal.* **1999**, *187*, 109–122.

47 Raybaud, P.; Hafner, J.; Kresse, H.; Toulhoat, H. *Phys. Rev. Lett.* **1998**, *80*, 1481–1484.

48 Raybaud, P.; Hafner, J.; Kresse, H.; Toulhoat, H. *Surf. Sci.* **1998**, *407*, 237–250.

49 Toulhoat, H.; Raybaud, P.; Kasztelan, S.; Kresse, H.; Hafner, J. *Catal Today.* **1999**, 50, 629–636.

50 Raybaud, P.; Hafner, J.; Kresse, G.; Kasztelan, S.; Toulhoat, H. *J. Catal.* **2000**, *189*, 129–146.

51 Cristol, S.; Paul, J. F.; Payen, E.; Bougeard, D.; Hutschka, F.; Hafner, J. Proc. 2nd Intern. Symposium on Hydrotreatment and Hydrocracking of Oil Fractions, ed. by B. Delmond, J. F. Froment, and P. Grange, Studies in Surf. Sci. and Catal. **1999**, 127, 327–334.

52 Cristol, S.; Paul, J. F.; Payen, E.; Bougeard, D.; Clemendot, F.; Hutschka, F. *J. Phys. Chem.*, **2000**, *104*, 11220–11229.

53 Travert, A.; Nakamura, H.; van Santen, R. A.; Cristol, S.; Payen, E. *J. Am. Chem. Soc.* **2002**, *124*, 7084–7095.

54 Paul, J. F.; Payen, E. *J. Phys. Chem. B* **2003**, *107*, 4057–4064.

55 $DMol^3$ is available as part of Material Studio. Accelrys Inc. San Diego. USA. **2002**.

56 Delley, B. *J. Chem. Phys.* **1990**, 92, 508–517.

57 Hahn, T. *International Tables for Crystallography*, Kluwer Academic Publishers. Boston. **1966**.

58 Espinosa, E.; Alkorta, I.; Elguero, J.; Molins, E. *J. Chem. Phys.* **2002**, *117*, 5529–5542.

59 Tang, T.-H.; Deretey, E.; Knak Jensen, S. J.; Csizmadia, I. G. *Eur. Phys. J. D* **2006**, *37*, 217–222.

60 Gatti, C.; May, E.; Destro, R.; Cargnoni, F. *J. Phys. Chem. A* **2002**, *106*, 2707–2720.

61 Aray, Y.; Rodriguez, J.; Coll, S.; Rodríguez-Arias, E.; Vega, D. *J. Phys. Chem. B* **2005**, *109*, 23564–23570.

62 Leboeuf, M.; Koster, M.; Jug, K.; Salahub, D. R. *J. Chem. Phys.* **1999**, *111*, 4893–4905.

63 Gadre, S. R.; Shirsat, R. N., *"Electrostatics of Atoms and molecules"*, Universities Press, Hyderabad, **2000**.

64 (a) Pingale, S. S.; Gadre, S. R.; Batolotti, L. J. *J. Phys. Chem. A* **1998**, *102*, 9987–9992; (b) Deshmukh, M. M.; Sastry, N. V.; Gadre, S. R. *J. Chem. Phys.* **2004**, *121*, 12402–12410.

65 (a) Pathak, R. K.; Gadre, S. R. *J. Chem. Phys.* **1990**, *93*, 1770; (b) Gadre, S. R.; Kulkarni, S. A.; Shrivastava, I. H. *J. Chem. Phys.* **1992**, *96*, 5253.

66 Gadre, S. R.; Pathak, R. K. *Proc. Ind. Acad. Sci. (Chem. Sci).* **1990**, *102*, 189.

67 (a) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem.* **1995**, *99*, 3084–3092; (b) Kornelak, P.; Michalak, A.; Najbar, M. *Catal. Today.* **2005**, *101*, 175–183.

68 Politzer, P.; Truhlar, D. G. *Chemical Applications of Atomic and Molecular*

Electrostatic Potentials, Plenum, New York, **1982**.

69 Murray, J. S.; Sen, K. D. *Molecular Electrostatic Potential: Concepts and Applications*, Elsevier, Amsterdam, **1996**.

70 Orozco, M.; Luque, F. J. in *Theoretical and Computational Chemistry Series (Ed: J. S. Murray, K. D. Sen)*, *Elsevier, Amsterdam*, **1996**, *3*, 181.

71 Aray, Y.; Marquez, M.; Rodríguez, J.; Coll, S.; Simón-Manso, Y.; Gonzalez, C.; Weitz, D. A. *J. Phys. Chem. B*. **2003**, *107*, 8946–8952.

72 (a) Aray, Y.; Marquez, M.; Rodríguez, J.; Coll, S.; Simón-Manso, Y.; Gonzalez, C.; Weitz, D. A. *J. Phys. Chem. B*. **2004**, *108*, 2418–2424; (b) Aray, Y.; Rodríguez, J.; Coll, S.; Gonzalez, C.; Marquez, M. *J. Phys. Chem. B*. **2004**, *108*, 18942–18948.

73 Tal, Y.; Bader, R. F. W.; Erkku, J. *Phys. Rev. A*, **21**, (1980), 1–11.

74 Keith, T. A.; Bader, R. F.; Aray, Y. *Int. J. Quantum Chem.* **1996**, *57*, 183–198.

75 Sjoberg, P.; Politzer, P. *J. Phys. Chem.* **1990**, *94*, 3959–3961.

Part III
Experimental Electron Densities and Biological Molecules

10
Interpretation of Experimental Electron Densities by Combination of the QTAMC and DFT

Vladimir G. Tsirelson

10.1
Introduction

Richard Bader summarized his longstanding studies dealing with development of the quantum theory of atoms in molecules and crystals (QTAMC) in his classic book published in 1990 [1]. He demonstrated that the ground-state electron density $\rho(\mathrm{r})$, its gradient vector field, $\nabla\rho(\mathrm{r})$, and the Laplacian field, $\nabla^2\rho(\mathrm{r})$, enable definition of bond paths, the lines of maximum electron density linking some of the nuclei, which can be identified at the equilibrium geometry, with the chemical bonds, and enable characterization of the type of these bonds in molecules and crystals. Originally, QTAMC was developed using electron density calculated from the wavefunctions. Later, it was demonstrated [2–5] that electron density derived from results from accurate X-ray, γ-ray, and synchrotron radiation diffraction experiments could also be analyzed in the same manner. Initial application of QTAMC to the experimental electron density of compounds with different types of chemical bond [6–11] showed that this function has a similar topology and the same set of critical points as quantum-mechanical ρ. Thus, the experimental electron density seems to be suitable for the QTAMC analysis of bonding in molecules and crystals, with electron density deduced from the wavefunctions. This approach is now widely used for exploration of experimental features of electron density; a thorough review of the results obtained from this type of work is available elsewhere [12–19].

In addition to electron density and its derivatives, Bader has also described the role of the positively-defined electronic kinetic energy density:

$$g(\mathbf{r}) = (\hbar^2/2m)\nabla_{\mathrm{r}}\nabla_{\mathrm{r}'}\gamma(\mathbf{r},\mathbf{r}')_{|\mathbf{r}=\mathbf{r}'} \tag{1}$$

and the potential energy density which is negative everywhere, defined by expression:

$$v(\mathbf{r}) = -\sum_{\mathrm{a}}[\mathrm{Z_a}e^2/(\mathbf{r}-\mathbf{R}_{\mathrm{a}})]\rho(\mathbf{r}) + e^2\int[\Gamma(\mathbf{r},\mathbf{r}_1)/(\mathbf{r}-\mathbf{r}_1)]\,\mathrm{d}\mathbf{r}_1 + \mathrm{V_n} \tag{2}$$

where $\gamma(\mathbf{r},\mathbf{r}')$ and $\Gamma(\mathbf{r},\mathbf{r}_1)$ are one-electron and two-electron density matrices, respectively, Z_a is the charge the nucleus a, e and m are the electronic charge and mass, respectively, and V_n is the nuclear repulsion. These functions are related by the local form of the virial theorem [20]:

$$2g(\mathbf{r}) + v(\mathbf{r}) = (\hbar^2/4m)\nabla^2\rho(\mathbf{r}) \tag{3}$$

Finally, the density of the total electronic energy is defined as:

$$h_e(\mathbf{r}) = g(\mathbf{r}) + v(\mathbf{r}) \tag{4}$$

Exploration of energy distributions in molecules by use of wavefunction-based calculations [21–26] has revealed that analysis of the local electronic energy is a direct approach to characterization of bonding in molecules and crystals. In particular, it facilitates recognition of the type of atomic interactions from the properties of bond critical points at $\mathbf{r}_b - h_e(\mathbf{r}_b) < 0$ and $g(\mathbf{r}_b)/\rho(\mathbf{r}_b) < 1$, and $\nabla^2\rho(\mathbf{r}_b) < 0$, are observed for shared-type atomic interactions whereas $h_e(\mathbf{r}_b) \lesseqgtr 0$, $g(\mathbf{r}_b)/\rho(\mathbf{r}_b) > 1$, and $\nabla^2\rho(\mathbf{r}_b) > 0$ are typical for intermediate and closed-shell interactions [16, 18, 23, 27].

Bader has also stressed [26, 28] that the potential energy density represents the field of the virial of the Ehrenfest force [29] acting on an electron at $\mathbf{r}$, the virial field $v(\mathbf{r})$. Irrespective of the type of atomic interaction, each bond path in the electron density at equilibrium geometry is homeomorphically mirrored by a virial path, a line of maximum negative potential energy density linking the same nuclei [24]. The presence of the bond paths and virial path provides, according to Ref. [26], an indicator of bonding atomic interaction. A network of the virial and bond paths defines a molecular graph, which is independent of the nuclear vibrations in a stable system.

It was, of course, a challenge to perform real-space energy analysis of bonding on the basis of the experimental electron density (ED). There is, fortunately, a theory which establishes the interconnection between the electron density and energy densities of different kinds. This is the density-functional theory (DFT) [30–40] which exploits ρ as a main variable and determines all the properties of atoms, molecules, and crystals in the ground electronic state [41]. Thus, DFT is a basis for quantitative characterization of bonding in terms of energy densities and other functions related to electron density. It is, therefore, attractive to combine the formalism of the DFT with experimental electron density to analyze the nature of atomic and molecular interactions in molecules and solids in terms of the local energies. This might, in principle, be done in two different ways. The exact functionals connecting ρ and the energy densities of electrons – the kinetic, potential, exchange and correlation densities – are, in general, unknown [41]. DFT methods therefore use either the Kohn and Sham orbital scheme [42] or approximate functionals with explicit (but nonunique) dependence of these functions on ρ and its derivatives [31]. The former approach might be achieved by use of an idempotent one-electron density matrix iteratively reconstructed from

the experimental electron density [12] or by use of Hartree–Fock calculations constrained to obtain the wavefunctions that reproduce experimental X-ray structure factors [43]. The latter approach is closer to the Hohenberg–Kohn formulation of the DFT [41] – it is orbital-free and enables avoidance of the variational determination of wavefunctions. This is the approach used in this chapter – our objective is to demonstrate here that it enables more comprehensive extraction of the information about chemical bonding contained in the experimental electron density.

10.2 Specificity of the Experimental Electron Density

First, we must clarify what is implied by "experimental" electron density. Experimental X-ray structure factors enable reconstruction of the electron density by means of Fourier series. Such density is strongly distorted by series truncation and is hardly suitable for quantitative topological analysis [12]. The electron density is therefore reconstructed from experiment by means of the multipole model in which the electron density of a molecule or crystal is presented as a sum of aspherical pseudoatomic densities, $\rho_{\mathrm{atom}}(\mathbf{r})$, each of which is expanded into a convergent series over the spherical harmonics or over their real combinations, $y_{lm\pm}$. There are a few multipole models, differing in insignificant details [12, 13]. In one of these, the Hansen–Coppens model [44], which we use in this work, the pseudoatomic density has the form:

$$\rho_{atom}(\mathbf{r}) = \rho_{core}(r) + P_v\kappa'^3\rho_{val}(\kappa' r) + \sum_{l=0}^{4}\kappa''^3 R_l(\kappa'' r)\sum_{m=-l}^{l} P_{lm\pm}y_{lm\pm}(\mathbf{r}/r) \qquad (5)$$

where ρ_{core} and ρ_{val} are the atomic core and valence electron densities, respectively, described by the wavefunctions of the free atoms, κ' and κ'' are atomic valence-shell contraction–expansion terms, and P_{val} and $P_{lm\pm}$ are the multipole electronic populations. The radial density functions have the exponential form $R_l(\mathrm{r}) \sim r^{n_l}\exp(-\kappa''\xi r)$, where n_l is related to the principal quantum number of the atom. The electronic populations of multipoles and the κ terms are determined by least-square fit to the experimental structure factors.

All the multipole models are rather flexible; it is, however, necessary to keep several points in mind. The exponential term, ξ, in the radial density functions $R_l(\mathrm{r})$ corresponds to the best model fit to the restricted number of low-angle X-ray structure factors and depends mainly on the electron density of the valence subshells. At the same time, the orbital exponential terms of the Hartree–Fock or Kohn–Sham basis functions are derived from the requirement of a minimum energy of a system; they are, therefore, more sensitive to the distribution of core electrons whose energies are higher. The interatomic density associated with interference of the atomic orbital is, moreover, not accounted for completely in the multipole model consisting of the atomic-like terms. The same is true for the part of the valence electron density localized near the nuclei. Further:

- the cusp condition is not necessary preserved during model fitting;
- the multipoles on different centers are not orthogonal (the Hirshfeld multipole model [45] is an exception); and
- the orthogonality of the core and valence s-type density functions of the same pseudoatom is not fulfilled in the Hansen–Coppens multipole model.

An attempt to correct the last shortcoming [46] failed.

Generalizing previous experience [11–19] we can conclude that X-ray diffraction experiments yield a quasi-static model of electron density extrapolated to infinite resolution, which is, typically, as precise around the bond critical point as $\sim$0.05 e Å^{-3}. The experimental error becomes larger closer to nuclei and increases with the atomic number in the vicinity of nuclei; therefore "internal" atomic regions ($R \approx 0.3$ Å) are normally excluded from consideration. Despite this, it is well-documented that experimentally derived electron density distributions have similar topology and the same set of the critical points as the corresponding quantum-mechanically derived densities [12–19].

Recent studies have shown the situation is not as encouraging for the Laplacian of the electron density [47–49]. For the closed-shell and intermediate atomic interactions [1], the Laplacian is restored from experiment in reasonable agreement with direct wavefunction calculations. In this instance contraction of the density toward the bond path is small, $\nabla^2\rho(\mathbf{r}_b) > 0$ and the atomic-like presentation of the electron density by the multipole model is quite reasonable. For shared atomic interaction ($\nabla^2\rho(\mathbf{r}_b) < 0$), the electron density curvatures perpendicular to the bond path, λ_1 and λ_2, are determined with good accuracy; the multipole model has, however, failed to correctly describe the electron-density curvature along the bond path, λ_3. As a result, the overall error in the experimental Laplacian $\nabla^2\rho(\mathbf{r}_b) = \lambda_1 + \lambda_2 + \lambda_3$ in this example can reach 50% [48].

Thus, the experimental electron density reconstructed with current multipole models can be regarded as an approximate homeomorphic image of the "true" ρ derived from the first principles, which is reasonably accurate for the closed-shell and intermediate atomic interactions and has significant quantitative uncertainty along the bond line for shared atomic interactions.

10.3 Approximate Electronic Energy Densities

10.3.1 Kinetic and Potential Energy Densities

The key problem in DFT is to express the kinetic, potential, and total electronic energy in terms of $\rho(\mathbf{r})$ [31, 32, 38]. The problem is the same for the energy densities. One way of solving this problem for kinetic energy density uses the fact

that the one-electron density matrix $\gamma(\mathbf{r}, \mathbf{r}')$, which defines the kinetic energy by use of Eq. (1), is related to the one-particle Green function by the inverse Laplace transform [38]. The gradient expansion of the Green function around the classical Thomas–Fermi approximation [50] leads to the following approximate expression for kinetic energy density:

$$g_{DFT}(\mathbf{r}) = (3\hbar^2/10m)(3\pi^2)^{2/3}\rho(\mathbf{r})^{5/3} + (\hbar^2/72m)[\nabla\rho(\mathbf{r})]^2/\rho(\mathrm{r}) + (\hbar^2/6m)\nabla^2\rho(\mathbf{r}) \quad (6)$$

Application of this formula, which is valid for smooth (but not necessarily small) variation of the electron density, to real systems implies the local homogeneity approximation [41] – the energy density at each point $\mathbf{r}$ is supposed to be the same as that of a homogeneous electron gas with the same electron density, which is equal to $\rho(\mathbf{r})$ everywhere. Note the role of Laplacian term in Eq. (6) [32, 51–60] – although it does not affect the average total and atomic energies its presence provides a description of the electronic shells and improves the local kinetic energy behavior of the valence electrons.

Consideration of the asymptotic properties of the kinetic energy density, Eq. (1), derived from the one-electron density matrix shows [20] that the long-range behavior of the approximate $g_{DFT}(\mathbf{r})$, Eq. (6), is physically acceptable [60]. In contrast, the function $g(\mathbf{r})$, Eq. (1), becomes $(1/2)Z^2\rho_i(\mathbf{R}_i)$ as $\mathbf{r} \to \mathbf{R}_i$, where $\rho_i(\mathbf{R}_i)$ is the value of the electron density at the positions of the nuclei, $\mathbf{R}_i$, while approximate $g_{DFT}(\mathbf{r})$, Eq. (6), becomes minus infinity as $\mathbf{r} \to \mathbf{R}_i$, because of the Laplacian term. The radius of the negative hole around the nuclei is maximum for the hydrogen atom (0.15 Å), is less than 0.02 Å for atoms with $Z \geq 11$ and reaches 0.005 Å for $Z = 36$ (Kr). This observation fits the 1/Z-dependency of this radius [55]. From the consideration given above it follows that the physically meaningless negative g_{DFT} regions are completely within the region of uncertainty of this function, because of experimental and model errors in ED; they should therefore be excluded from the discussion during interpretation of the approximate $g_{DFT}(\mathbf{r})$. In other points of the position space, $g_{DFT}(\mathbf{r})$, Eq. (6), is quite close to the quantum mechanical $g(\mathbf{r})$, Eq. (1) [52, 61]. Thus, the use of the experimental electron density and its derivatives to determine the kinetic energy density has a physical basis.

Approximation Eq. (6) opens the way to determination of the potential energy density from X-ray experiments. It has been postulated [62, 63] that the model electron density derived by the fit to experimental structure factors does obey the local virial theorem, Eq. (3), the same as the quantum-mechanical $\rho(\mathbf{r})$ does. Then, using $g_{DFT}(\mathbf{r})$, Eq. (6), and $\nabla^2\rho(\mathbf{r})$, it is possible to obtain the potential energy density by use of Eq. (3) and calculate electronic energy density $h_e(\mathbf{r})$, Eq. (4), from the experimental ED. This approximation is nonevident; subsequent studies [61, 64, 65] have shown, however, that $g_{DFT}(\mathbf{r})$, Eq. (6), calculated from the experimental electron density leads to physically reasonable (negative everywhere) potential energy density.

After other work [58, 62, 63, 66–68], the approach summarized above became a popular tool for determination of the energy characteristics at the bond critical points of crystalline systems [14–19, 67–71]. It is, however, necessary to mention that the above-mentioned inability of the current multipole models to correctly describe the curvature of the experimental electron density along the bond path for shared atomic interaction influences all existing QTAMC bonding descriptors which contain the Laplacian term. Fortunately, electron density for the closed-shell and intermediate atomic interactions is reconstructed correctly and corresponding kinetic energy density approximated by use of Eq. (6) behaves properly in the internuclear space [72]. Equations (6), (2), and (4) are therefore completely applicable to systems with closed shells and intermediate atomic interactions.

To illustrate the actual situation with the applicability of the gradient expansion, Eq. (6), to the determination of the energy densities, let us consider these functions for urea, $CO(NH_2)_2$. The electron density in a single urea molecule removed from a crystal was reconstructed from data from two diffraction experiments – an X-ray four-circle diffractometer experiment at 148 K [73] and an X-ray synchrotron experiment at 123 K [74]. Multipole data derived from X-ray and synchrotron results were taken from Refs [73] and [75]. The approximate kinetic energy density, $g_{EXP/DFT}(\mathbf{r})$, and the potential energy density, $v_{EXP/DFT}(\mathbf{r})$, were calculated from the multipole data by use of Eq. (3) and Eq. (6). We then calculated the wavefunction for the urea molecule by the Hartree–Fock method in the 6-311G** basis set using the PC version [76] of the GAMESS software [77]. The optimized molecule geometry was taken from Ref. [61]. First, the gradient expansion Eq. (6) was used to calculate the kinetic energy density $g_{HF/DFT}(\mathbf{r})$ from the wavefunction and the local potential energy $v_{HF/DFT}(\mathbf{r})$ was calculated by means of the local virial theorem, Eq. (3), which is valid in the Hartree–Fock theory. Second, the same functions were also calculated directly from the Hartree–Fock wavefunctions using the AIMPAC software suite [78]. The later functions will subsequently be referred to as $g_{HF}(\mathbf{r})$ and $v_{HF}(\mathbf{r})$, respectively.

By comparing electron densities derived for urea by different methods (Fig. 10.1), we see that both experimental EDs are in very reasonable agreement, excluding the vicinity of the hydrogen nuclei (we ignore small distortions of the density present on the periphery of molecules because of the effect of the crystalline environment). More complete and accurate synchrotron-measured electron density is also in quantitative agreement with the Hartree–Fock electron density practically everywhere in the position space. Functions $g_{EXP/DFT}(\mathbf{r})$ derived from the data from both experiments (Fig. 10.2) are in remarkable mutual quantitative agreement (again excluding the vicinity of the hydrogen nuclei); they differ, however, from results from direct Hartree–Fork calculation, $g_{HF}(\mathbf{r})$ (note that the geometry of the urea molecule in a crystal depends on the temperature of the experiments and is different from that in the free state). A similar difference is observed between $g_{HF}(\mathbf{r})$ and $g_{HF/DFT}(\mathbf{r})$. The approximate function $g_{HF/DFT}(\mathbf{r})$ and both experimental kinetic energy densities reconstructed in the same manner via the multipole model are in close agreement.

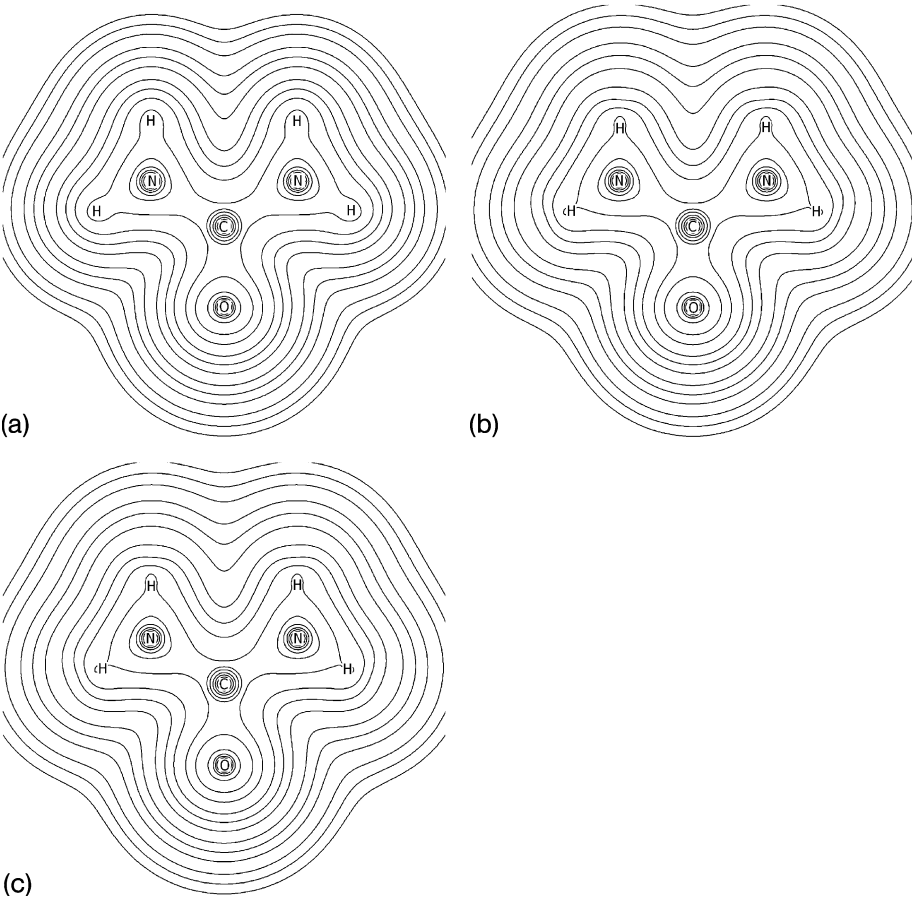


Fig. 10.1 Electron densities of urea derived by different methods: (a) reconstructed from the synchrotron diffraction experiment [74], (b) reconstructed from the X-ray diffraction experiment [73], (c) calculated theoretically by use of the Hartree–Fock /6-311G** method. Line intervals are $(2, 4, 8) \times 10^n$ e Å^{-3} $(-2 \leq n \leq 2)$.

Both functions $v_{EXP/DFT}(\mathbf{r})$ derived from the experimental data (Fig. 10.3) are also in close agreement; they differ from $v_{HF}(\mathbf{r})$ only in secondary details.

Let us now compare the experimental and Hartree–Fock energy densities for ionic LiF crystal. Multipole-model data were taken from Ref. [79]. Hartree–Fock (6-311G*) calculations were performed for the cubic-like cluster $Li_{14}F^{+}{}_{13}$ of optimized geometry [80] surrounding the central fluorine ion to simulate a crystal. Experimental and theoretical electron densities of LiF are in evident agreement (Figs 10.4a and 10.4b). Comparison of the approximate DFT-based local energy densities $g_{EXP/DFT}(\mathbf{r})$ and $v_{EXP/DFT}(\mathbf{r})$ with results from direct Hartree–Fock calcu-

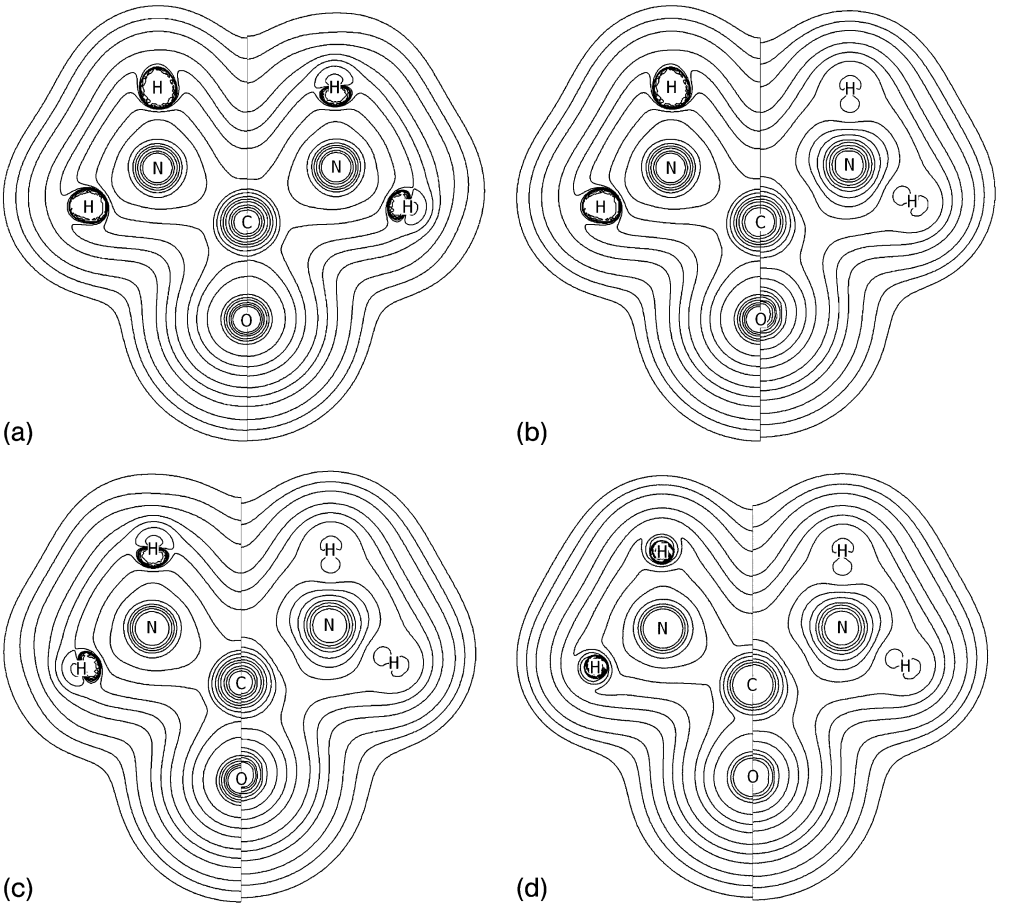


Fig. 10.2 Distributions of the kinetic energy in the urea molecule. Left and right parts of each figure represent the local kinetic energies obtained from different sources: (a) left, synchrotron data; right, X-ray diffraction data; (b) left, synchrotron data; right, Hartree–Fock result; (c) left, X-ray diffraction data; right Hartree–Fock result; (d) left, calculation by use of Eq. (6) and Hartree–Fock wavefunctions; right, Hartree–Fock result. Line intervals are $(2, 4, 8) \times 10^n$ atomic units ($-2 \leq n \leq 2$). The geometry of the free molecule and of the molecule in the crystal are different, which is why the maps were merged in such a way that the positions of the C atoms coincided.

lation were also in quantitative agreement everywhere in space with maximum deviation of ~0.01 a.u. in the low-density region of $v_{EXP/DFT}(\mathbf{r})$ around the center of the (100) plane of the cubic unit cell.

Thus the quantitative agreement between Hartree–Fock and experimentally modeled energy densities is reached for the closed-shell (and intermediate [81]) atomic interactions with the leading atomic-like electron density contributions. The failure of the multipole model to correctly describe the electron-density cur-

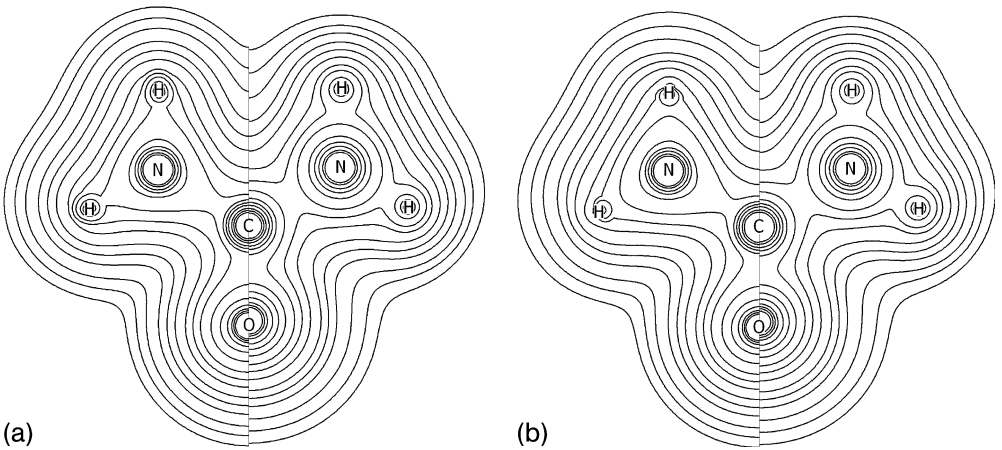


Fig. 10.3 Distribution of the potential energy in the urea molecule. Left and right parts of each figure represent the local potential energies obtained from different sources: (a) left, synchrotron data; right, Hartree–Fock result; (b) left, X-ray diffraction data; right, Hartree–Fock result. Line intervals are $(2, 4, 8) \times 10^n$ atomic units $(-2 \leq n \leq 2)$.

vature along the bond path for shared atomic interactions is, however, a major source of distortion of the energy distribution around the bond critical points. At the same time it is essential that the local energy functions $g_{EXP/\mathrm{DFT}}$ and $v_{EXP/\mathrm{DFT}}$ have the same topology as the corresponding Hartree–Fock functions. We expect that use of more flexible radial functions in the multipole model, as it discussed elsewhere [47, 48], will remove this defect.

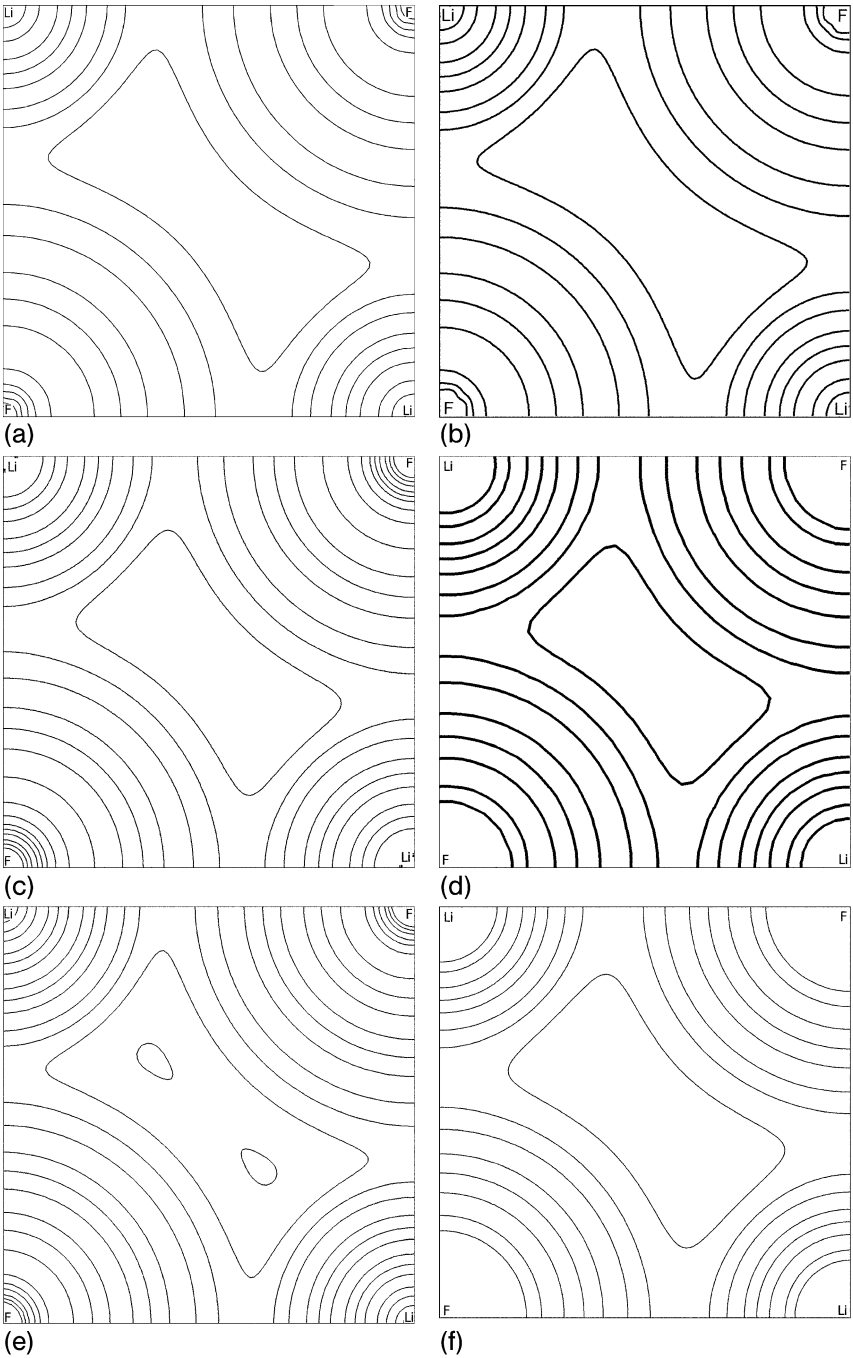
This shortcoming of the model density is less vital for shared atomic interactions (because of partial compensation of contributions) if, in accordance with Ref. [21], one calculates the difference functions:

$$\delta g(\mathbf{r}) = g(\mathbf{r}) - g_{\mathrm{pro}}(\mathbf{r}) \tag{7}$$

and

$$\delta v(\mathbf{r}) = v(\mathbf{r}) - v_{\mathrm{pro}}(\mathbf{r}) \tag{8}$$

(the suffix ''pro'' denotes functions calculated for a procrystal – a hypothetical system consisting of spherical noninteracting atoms placed in the same positions as real atoms). These difference functions reveal, at the semiquantitative level, the changes in corresponding energy densities caused by formation of a crystal from the atoms. Figs 10.5 and 10.6 depicting the functions $\delta g(\mathbf{r})$ and $\delta v(\mathbf{r})$ for crystalline urea and LiF, respectively, explicitly demonstrate the difference between the covalent and ionic bonding mechanisms as reflected in the energy distribution in


Li
F
F
Li
(a)
Li
F
F
Li
(b)
Li
F
F
Li
(c)
Li
F
F
Li
(d)
Li
F
F
Li
(e)
Li
F
F
Li
(f)

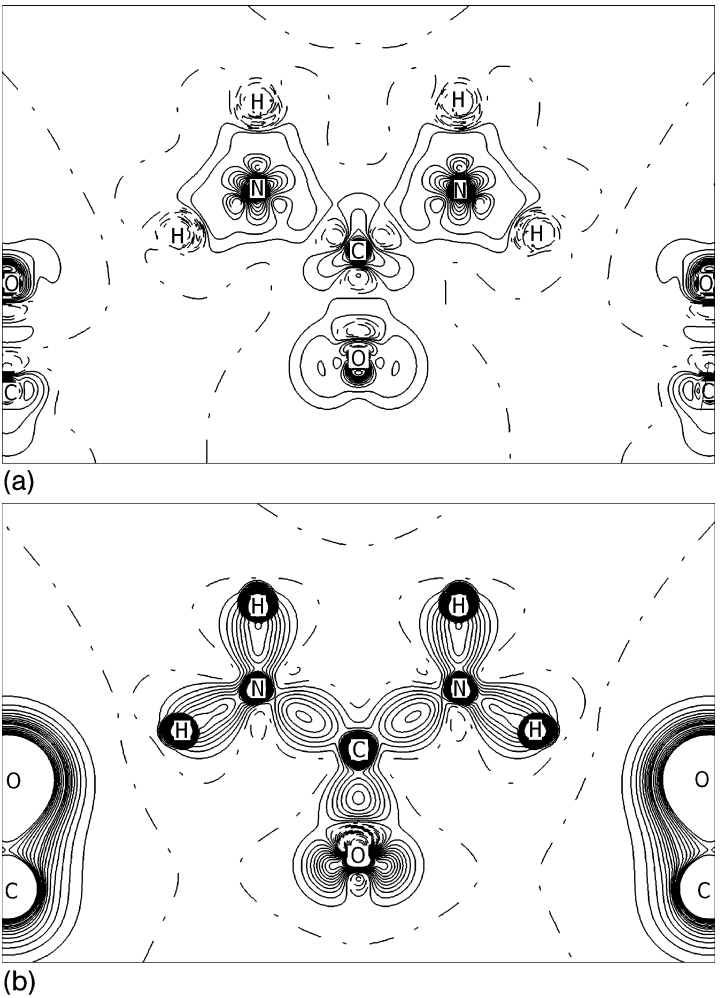


Fig. 10.5 Crystalline urea – the difference functions $\delta g(\mathbf{r}) = g(\mathbf{r}) - g_{pro}(\mathbf{r})$ (a) and $\delta\nu(\mathbf{r}) = \nu(\mathbf{r}) - \nu_{pro}(\mathbf{r})$ (b) characterizing changes in the kinetic and potential energy densities, correspondingly, caused by formation of a crystal from the atoms. Line intervals are 0.05 a.u. Solid lines correspond to excessive (positive) kinetic energy density and (negative) potential energy density.

Fig. 10.4 Distributions of electron density and kinetic and potential energy in the (100) plane of crystalline LiF: (a) experimental electron density, (b) Hartree–Fock (cluster) electron density, (c) experimental kinetic energy density, (d) Hartree–Fock kinetic energy density, (e) experimental potential energy density, (f) Hartree–Fock potential energy density. Line intervals are $\pm(2, 4, 8) \times 10^n$ a.u. $(-2 \leq n \leq 2)$.

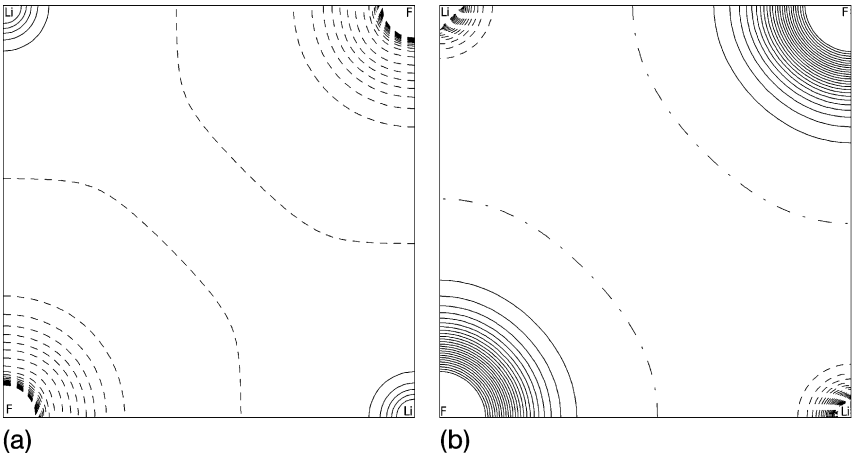


Fig. 10.6 Crystalline LiF: the difference functions $\delta g(\mathbf{r}) = g(\mathbf{r}) - g_{pro}(\mathbf{r})$ (a) and $\delta v(\mathbf{r}) = v(\mathbf{r}) - v_{pro}(\mathbf{r})$ (b). Line intervals are 0.05 a.u. Solid lines correspond to excessive (positive) kinetic energy density and (negative) potential energy density.

position space. In urea, the kinetic energy density increases strongly in the internal electronic shells of nonhydrogen atoms and only slightly (and irregularly) in the intramolecular bonds and electron lone-pair regions. In contrast, the potential energy density regularly increases in the intramolecular bonds and in the electron lone pairs (and within the cores of nonhydrogen atoms). Distributions of both functions in the hydrogen bond areas are close to the superimposition of the free atoms. Thus, distributions $\delta g(\mathbf{r})$ and $\delta v(\mathbf{r})$ show the stabilizing enhancement in the potential energy along the bond lines and in the electron lone pairs of the oxygen atom and details of destabilizing contributions of the local kinetic energy during formation of crystalline urea from atoms.

The distribution of functions $\delta g(\mathbf{r})$ and $\delta v(\mathbf{r})$ in the ionic LiF crystal (Fig. 10.6) is dramatically different. Formation of this crystal from the neutral atoms is accompanied by concentration of the kinetic energy in the Li atom basins and more pronounced enhancement in the (negative) local potential energy within the F atomic basins. This reveals the stabilizing role of the anions during LiF crystal formation. Note that the areas of the energy concentration/depletion in LiF are close to spherical.

Consideration of the functions $\delta g(\mathbf{r})$ and $\delta v(\mathbf{r})$ for some organic compounds and cubic perovskite $SrTiO_3$ [82, 83] led to similar conclusions. It was also noted [82] that partial covalence of the Ti–O bond in $SrTiO_3$ manifests itself in the noticeable dipole-type $\delta v(\mathbf{r})$ distribution and small excessive dipole-type $\delta g(\mathbf{r})$ distribution around the O atoms directed to the Ti atoms; it explicitly exhibits the polar bonding contribution in the Ti–O closed-shell interaction. Thus, the energy distribution features can help in explicit topological electron density characterization of the polar shared atomic interactions (or partially ionic bonds) and provide a new insight into bonding mechanisms from data of the X-ray diffraction experiment.

The energy distributions depicted in Figs. 10.1–10.6 show the importance of consideration of all the position space in studies of bonding in molecular systems. No significant details are seen in the position–space energy distributions in ionic LiF crystal, therefore information about atomic interactions, which is concentrated in the bond critical points, is sufficient for bond description. In contrasts, many energy-density features reflecting the bonding mechanism in urea (for example, the electron lone pairs) cannot be taken in account if only the bond critical points are considered.

It is well known that the rapid variation of the electron density in the vicinity of the nuclei and its slow variation in the valence electron shells makes it difficult to find a functional for kinetic energy density which provides a good description everywhere in the position space [32, 60]. We have tested a few other approximations for describing the kinetic energy density via the electron density and its derivatives; our observations can be briefly summarized as follows. Inclusion of the 4th-order correction to the gradient expansion Eq. (6) [84] did not result in discernible improvement. The Lee–Lee–Parr formula [85] (plus the Laplacian term) yielded $g_{DFT}(\mathbf{r})$ in quantitative agreement with the gradient expansion. Zero-order presentation of the Green function in the mean-path approximation using the Feynman path–integral method [54] yields an expression differing from Eq. (6) by the numerical coefficients in front of the gradient corrections; this resulted in physically meaningless areas with $v_{DFT}(\mathbf{r}) > 0$ in the low-electron-density regions, for example the periphery of free molecules or the centers of the faces of a cubic unit cell of LiF. Kinetic energy density derived using the virial theorem relationships of density-functional theory [57] seemed to depend on the origin of the coordinate system; this makes its application to molecules and crystals difficult. And, finally, the Weizsacker approximation [86], which corresponds to rapidly oscillating density (as happens in atomic cores), has failed to describe the middle-bond areas.

It would be interesting to test the hybrid orbital-free energy functionals [87] to search for the better energy-density description.

10.3.2
Exchange and Correlation Energy Densities

In DFT, the exchange-correlation energy, E_{xc}, describes a contribution of nonclassical electron–electron interaction to total energy [88]. E_{xc} is usually decomposed into the exchange, E_x, and the correlation, E_c, parts [31]; correspondingly, the exchange, ε_x, and correlation, ε_c, energy densities are often discussed. These functions are not uniquely defined (they can, for example, be altered by addition of any functional of ρ that integrates to zero over the density, or by coordinate transformation [88]); they also have a different physical content in different versions of DFT [88]. The densities ε_x and ε_c nevertheless play an important role in DFT; they also extend naturally a set of the functions, which are considered in QTAMC.

Many approximations have been developed which enable expression of ε_x and

ε_c via $\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$, and $\nabla^2\rho(\mathbf{r})$ (reviewed elsewhere [38]). We consider here the gradient-corrected exchange energy density by Becke [89], which is expressed for spin-unpolarized (i.e. the closed-shell) systems as:

$$\varepsilon_x = \varepsilon_{x,\text{uniform}} - (1/2)^{1/3}\frac{bX^2}{1+6bX-\sinh^{-1}(X)}\rho(\mathbf{r})^{4/3} \tag{9}$$

where:

$$\varepsilon_{x,uniform} = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho^{4/3}(\mathbf{r}), \quad b = 0.0042, \quad X = \frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma(\mathbf{r})^{4/3}}$$

$\varepsilon_{x,\text{uniform}}$ is the Dirac–Slater exchange density for a uniform electron gas; the value of $(1/2)\rho(\mathrm{r})$ is assigned for the spin density $\rho_\sigma(\mathbf{r})$. We consider also the gradient-corrected correlation energy density $\varepsilon_c(\mathbf{r})$ by Perdew–Burke–Ernzerhof [90]:

$$\varepsilon_c(\mathbf{r}) = [e_c(\mathbf{r}) + H(\mathbf{r})]\rho(\mathbf{r}) \tag{10}$$

where $e_c(\mathbf{r})$ is correlation energy per electron for a uniform electron gas [91] and:

$$H(\mathbf{r}) = \gamma\ln\left[1+\frac{\beta}{\gamma}t^2(\mathbf{r})\left(\frac{1+A(\mathbf{r})t^2(\mathbf{r})}{1+A(\mathbf{r})t^2(\mathbf{r})+A^2(\mathbf{r})t^4(\mathbf{r})}\right)\right],$$

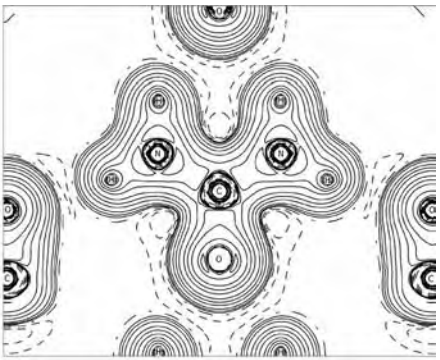
$$A(\mathbf{r}) = \frac{\beta}{\gamma\exp\left(-\frac{e_c(\mathbf{r})}{\gamma}\right)-1}, \quad t(\mathbf{r}) = |\nabla\rho(\mathbf{r})|/[2k_s\rho(\mathbf{r})], \quad k_s(\mathbf{r}) = 2\left[\frac{3}{\pi}\rho(\mathbf{r})\right]^{1/6},$$

($\beta = 0.066725$, $\gamma = 0.031091$). Approximations Eq. (9) and Eq. (10) are popular in DFT applications [38, 40, 88] and we use them to explicitly reveal regions of potential energy reduction in crystalline urea, $CO(NH_2)_2$, caused by exchange between electrons of the same spin and spin-independent electron correlation.

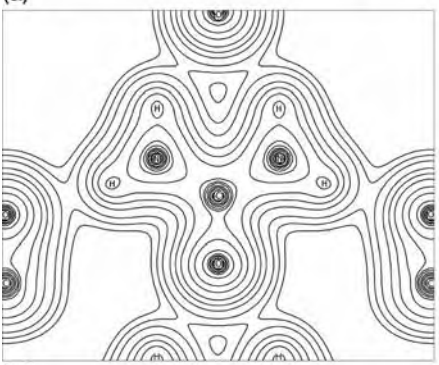
The structure of urea (space group $P\bar{4}2_1m$, Z = 2 (2 mm)) is characterized by ribbons of hydrogen-bonded molecules arranged head-to-tail along the c axis (Fig. 10.7a). The plane of each ribbon is perpendicular to adjacent ribbons directed oppositely along the c axis. The oxygen atom of a carbonyl group is

Fig. 10.7 Structure of crystalline urea (a) and distributions of the electronic energy density $h_e(\mathbf{r})$ (b), the gradient-corrected exchange energy density ε_x (c), Dirac–Slater exchange energy density (d), gradient-corrected correlation energy density (e), and correlation energy density for a uniform electron gas (f). Line intervals are $\pm(2,4,8)\times10^n$ a.u. ($-3 \le n \le 3$); negative values are solid.

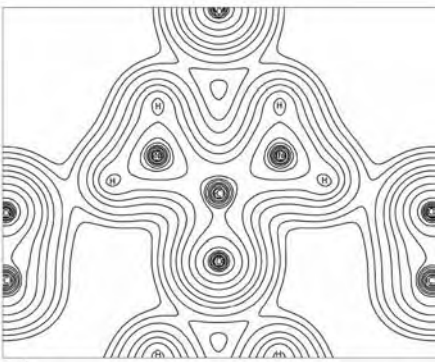
(a)



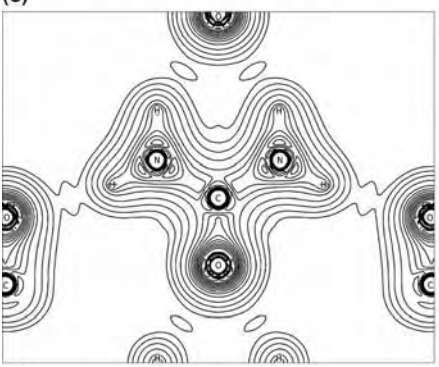
(b)



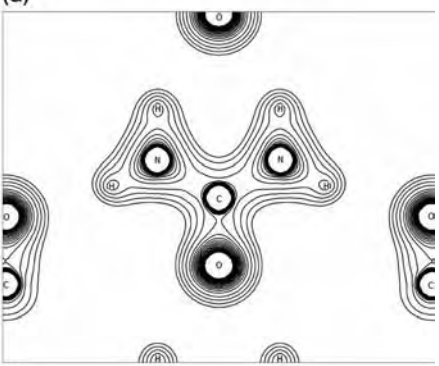
(c)



(d)



(e)



(f)

H-bonded to the neighboring molecule in the same ribbon and, simultaneously, is involved in H-bonds with two adjusted ribbons. Distributions of the exchange and correlation energy densities in urea are shown in Fig. 10.7, with total electronic density $h_e(\mathbf{r})$ Eq. (4). The energy density $h_e(\mathbf{r})$ (Fig. 10.7b) achieves its lowest negative values on the intramolecular bond lines, revealing areas of concentration associated with bonding interactions and lone pairs. (These features cannot be attributed to the presence of the Laplacian term in Eq. (6) – the Hartree–Fock electronic energy density has the same features.) Simultaneously, alternating negative minima and positive maxima observed in the vicinity of nuclei have the shell structure of bonded O, N, and C atoms. At the same time, $h_e(\mathbf{r})$ is slightly positive in the intermolecular hydrogen bonds in urea, where kinetic energy dominates. The function $h_e(\mathbf{r})$ enables hydrogen bonds of different length to be distinguished: the longer H-bond ($d_{O\ldots H2} = 2.064(2)$ Å) is characterizing by a minimum value of $h_{e,\min} = +0.00179$ a.u. whereas the shorter H-bond ($d_{O\ldots H1} = 2.007(2)$ Å) has the less positive value $h_{e,\min} = +0.00139$ a.u. Similar behavior of $h_e(\mathbf{r})$ in crystalline urea (148 K data) and α-oxalic acid dihydrate, $C_2H_2O_4{\cdot}H_2O$, has been reported elsewhere [61, 92]. This also agrees with a general observation made for weakly bounded molecular systems [25] and weak and intermediate hydrogen bonds [93].

A map of the local exchange energy ε_x (Eq. 9), depicted in the same plane (Fig. 10.7c), in addition to deep energy wells in the vicinity of the nuclei, also has negative exchange-energy density bridges between bounded atoms, which contributes to the potential energy reduction during the crystal formation. The magnitudes of exchange contributions reflect the features of a bond, e.g. the shorter H-bond in urea is characterized by a lower value of $\varepsilon_x = -0.005$ a.u. than its longer counterpart (-0.004 a.u.).

Function ε_x does not reproduce explicitly such typical bonding features as the bond charge and electron lone-pair concentrations or core density alternations. Although the last of these may be made evident by calculating $\nabla^2[-\varepsilon_x(\mathbf{r})]$, there is no need to do this because the Laplacian of electron density itself provides us with this information.

It is worth noting that the gradient-corrected exchange density Eq. (9) is quite similar to the Dirac–Slater exchange density for a uniform electron gas, $\varepsilon_{x,\,uniform}$, depicted in Fig. 10.7d. The main discrepancy is observed in the vicinity of nuclei whereas on the bond lines it is only 0.002 a.u.

A map of the local gradient-corrected correlation energy, ε_c, (Fig. 10.7e) shows the contribution of the electron correlation to the total electronic energy, which is independent of the electron spin. It also reveals the energy wells in the vicinity of the nuclei and the negative energy density bridges between bounded atoms, indicative of reduced potential energy. Simultaneously, in contrast with the function ε_x, the gradient-corrected correlation energy density ε_c has the typical bonding features of bond charge and lone-pair electron concentrations. These features are absent in the map of correlation energy for a uniform electron gas $\varepsilon_{c,\,uniform} = e_c(\mathbf{r})\rho(\mathbf{r})$ computed in accordance with Ref. [91] (Fig. 10.7f). Similar results were

obtained early for molecules [94]. Thus, analysis of the correlation energy density reconstructed from the experimental electron density in the gradient-corrected approximation can provide precise details of the bonding mechanisms.

10.4 The Integrated Energy Quantities

The average molecular energy calculated by use of the variational principle using the kinetic energy approximation, Eq. (6), is only qualitatively close to the experimental value [31, 60]. At the same time, computation of approximate g_{DFT}, Eq. (6), with the Hartree–Fock wavefunctions leads to average kinetic energy differing from the Hartree–Fock energy by approximately 1%. It is, therefore, worth determining the integrated energy-related quantities from the experimental ED combined with Eq. (6), especially because experimental ρ is close to the theoretical value.

According to Bader [1], the position space of a molecule or crystal may be divided into atomic basins separated by surfaces satisfying the zero-flux condition:

$$\nabla\rho(\mathbf{r}) \cdot \mathrm{n}(\mathbf{r}) = 0, \quad \forall \mathbf{r} \in \mathrm{S_i}(\mathbf{r}) \tag{11}$$

These basins are identified with bonded atoms (pseudoatoms). An integral of any property, A(r), over the volume of an atom i, Ω_i:

$$\langle \mathrm{A_i} \rangle = \int_{\Omega_i} \mathrm{A(r)}\, \mathrm{dV} \tag{12}$$

yields an average value of the property. These quantities are uniquely determined because $\int_{\Omega_i} \nabla^2\rho(\mathbf{r})\, \mathrm{dV} = 0$. The sum of atomic contributions thus obtained yields the value of the property for a whole system and for the functional atomic groups, bonded molecules, and elementary cells in a crystal.

Atomic components of electronic energy, H_e, for H_2O (in α-oxalic acid dihydrate, $C_2H_2O_4{\cdot}H_2O$), NH_3, and Cl_2 molecules removed from a crystal are given in Table 10.1. The integral $\int_{\Omega_i} \nabla^2\rho(\mathbf{r})\, \mathrm{dV}$ over each of the atomic basins was less than $\pm 10^{-3}$ a.u., thus the latter number is an estimate of the integration error. After summing, atomic contributions yield the electronic energy of a molecule. Quantum chemistry-calculated energy values H_e for free molecules (Table 10.1) are in reasonable agreement with "experimental" energies, despite the slight distortion of the experimental electron densities by the parent crystal environment. The largest discrepancy of 1.7% obtained for the Cl_2 molecule may be attributed to the relatively low accuracy of the corresponding X-ray diffraction experiment. We note, however, that even with the perfect X-ray diffraction data energies of in-

Table 10.1 Atomic contributions to the electronic energy, H_e, exchange energy, E_x, and correlation energy, E_c, calculated from the model electron densities for water (in α-oxalic acid dihydrate), ammonia, and chlorine molecules (all the molecules removed from a crystal). The total molecular values calculated by summing atomic contributions are also listed; the corresponding nonempirically calculated values are given in parentheses. All values are given in atomic units.

Molecule/atom	H_e	E_x	E_c
O	−75.743	−8.759	−0.322
H_2O H	−0.298	−0.162	−0.011
H	−0.296	−0.160	−0.011
Total	−76.337	−9.081	−0.344
	(−76.171 [100])	(−8.946 [101])	(−0.371 [102])
NH_3 N	−55.300	−7.129	−0.290
H	−0.301	−0.171	−0.012
Total	−56.204	−7.642	−0.326
	(−56.326 [100])	(−7.670 [101])	(−0.340 [102])
Cl_2 Cl	−451.517	−26.887	−0.667
Total	−903.034	−53.774	−1.334
	(−918.892 [100])	(−55.094 [101])	(−1.380 [102])

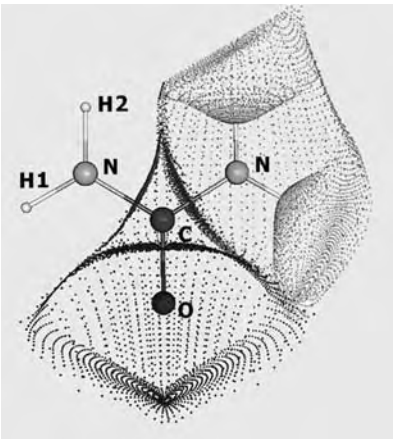
termolecular interaction or cohesion can hardly be obtained in this way, because their typical values are comparable with experimental uncertainty.

The atomic integrated values of the exchange energy, E_x, the correlation energy, E_c, and the corresponding molecular values are also given in Table 10.1. A maximum E_x discrepancy of 2.5% is observed for Cl_2 molecule, whereas a maximum E_c discrepancy of 7.5% is found for H_2O molecule. The last disagreement is partially the result of the different correlation energy definitions in quantum chemistry and DFT [88]. Overall agreement between experimentally derived and theoretically computed exchange and correlation energies is, nevertheless, reasonable. The same conclusion is valid for the crystalline wea (Table 10.2).

10.5 Concluding Remarks

Our approach consists in the joint use of the QTAMC and DFT to treat the electron density and density-dependent functions obtained with the multipole model is fitted to accurate X-ray diffraction data. Despite current problems with adequate

Table 10.2 Atomic volumes, Ω_i, restricted by zero-flux surfaces shown in the picture below, atomic electron populations, P_i, and the integrated atomic electronic, $H_{e,i}$, exchange, $E_{x,i}$, and correlation, $E_{x,i}$, energies of the urea molecule computed by integration of corresponding densities derived from the X-ray synchrotron data at 123 K (see text). Only values for symmetry-independent atoms are given.



Atom	Ω_i, Å^3	P_i, e	$H_{e,i}$, a.u.[a]	$E_{x,i}$, a.u.[a]	$E_{c,i}$, a.u.[a]
C	4.818	4.27(2)	−37.108(1)	−4.595(1)	−0.143(1)
O	16.486	9.19(2)	−75.498(1)	−7.043(1)	−0.324(1)
N	17.985	7.96(2)	−55.502(1)	−8.715(1)	−0.281(1)
H1	3.847	0.65(2)	−0.446(1)	−0.237(1)	−0.016(1)
H2	3.923	0.66(2)	−0.450(1)	−0.238(1)	−0.016(1)

[a] The sum of the atomic energy contributions over a single molecule is −225.402(1) a.u. The molecular Hartree–Fock/6-311G** energy is −224.046 a.u., whereas the B3LYP/6-311G** method yields −225.401 a.u.

description of the electron density for shared atomic interactions (which can be overcome by use of a more sophisticated multipole model), the level of unification of the theoretical and experimental methods achieved substantially increases the amount of information derivable directly from X-ray diffraction data. It reveals how the total electronic energy and its different components are distributed over a molecule and crystal and provides a real-space insight into the energetics of the chemical bond. We also hope that analysis of the exchange and correlation energy densities derived in different approximations from experimental electron density may be useful for improving existing DFT functionals.

It is worth mentioning that we did not consider all aspects of the aforesaid approach. Approximate determination of the electron localization function, localized-orbital locator, the local (and integrated) internal temperature of an electron gas, and associated entropy from the experimental electron density and its derivatives is reported elsewhere [92, 95, 96]; the kinetic energy density was described in these works by use of Eq. (6). Dirac–Slater exchange potential and one-electron potential determined using experimental electron density are de-

scribed elsewhere [27, 97]. Bonding analysis in terms of the critical points properties is also given elsewhere [18].

Calculation of all the energy characteristics from the electron density in this work was performed with the WinXPRO program [97, 98] using the relativistic many-configuration wavefunctions from Ref. [99].

Acknowledgments

I would like to express my deep gratitude to Richard Bader for longstanding scientific collaboration and personal friendship. The author thanks Dr A.I. Stash for collaboration and Professor W. Scherer and Dr H. Birkedal for making available the X-ray experimental data for α-oxalic acid dihydrate and urea. This work is supported by the Russian Ministry for Education (grant 2.1.1.5051) and the Russian Foundation for Basic Research (grant 04.03.33053).

References

1 R.F.W. Bader, *Atoms in Molecules – A Quantum Theory*. Oxford University Press, Oxford, **1990**.

2 V.A. Streltsov, V.G. Tsirelson, ''Laplacian of electron density new approach to the chemical bond analysis based on X-ray diffraction data''. *IV All-Union Conference on Inorganic and Coordination-Compound Crystal Chemistry*. Buchara. Abstracts, **1986**, p. 30.

3 M. Kapphahn, V.G. Tsirelson, R.P. Ozerov, ''Quantum-topological approach to the X-ray diffraction result interpretation''. *Portugal Phys.*, **1988**, *19*, 213–216.

4 N.N. Lobanov, E.L. Belokoneva, V.G. Tsirelson,'' The chemical bond in Ca–Nb–Ga garnet from X-ray diffraction data''. *Russ. J. Inorg. Chem.*, **1988**, *32*, 1740–1744.

5 M. Kapphahn, V.G. Tsirelson, & R.P. Ozerov, ''Quantum-topological interpretation of the X-ray diffraction data C_2H_6 molecule in a crystal''. *Doklady Acad. Nauk SSSR*, **1989**, *303*, 1025–1028.

6 M.Yu. Antipin, A.V. Polyakov, V.G. Tsirelson, M. Kapphahn, V.V. Grushin, Yu.T. Struchkov, ''Chemical bond in 9-asido-m-carborane''. *Organomet. Chem. USSR*, **1990**, *3*, 421–429.

7 V.G. Tsirelson, E.L. Belokoneva, V.A. Streltsov, M.Yu. Antipin, ''Electron density and exchange interaction in magnetically ordered crystals''. *X Sagamore Conference on Charge, Spin and Momentum Densities*. Abstracts. Konstanz, **1991**, 141.

8 V.G. Tsirelson, R.P. Ozerov, ''Electron density and chemical bond in organic compounds from the X-ray diffraction data''. *J. Mol. Struct. (THEOCHEM)*, **1992**, *255*, 335–392.

9 C. Gatti, R. Bianchi, R. Destro, F. Merati, ''Experimental vs. theoretical topological properties of charge density distributions. An application to the L-alanine molecule studied by X-ray diffraction at 23 K''. *J. Mol. Struct. (THEOCHEM)*, **1992**, *255*, 409–433.

10 V.G. Tsirelson, ''Topological characterization of electron density from X-ray diffraction data''. *XI Conference on Charge, Spin and Momentum Densities*. Abstracts. Bråst, **1994**, p. 252.

11 V.G. Tsirelson, P.F. Zou, T.-H. Tang, R.F.W. Bader, ''Topological definition of crystal structure determination of the bonded interactions in solid molecular chlorine''. *Acta Cryst.*, **1995**, *A51*, 143–153.

12 V.G. Tsirelson, R.P. Ozerov, *Electron Density and Bonding in Crystals*, Bristol and Philadelphia Inst. of Physics Publ., **1996**.

13 P. Coppens, *X-Ray Charge Densities and Chemical Bonding*. International Union of Crystallography, Oxford University Press, **1997**.

14 P.L.A. Popelier, F.M. Aicken, S.E. O'Brien, ''Atoms in Molecules''. In *Specialist Periodical Report, Chemical Modeling Applications and Theory* (Ed. Hinchliffe), A. Vol. 1. The Royal Society of Chemistry, **2000**, 143–198.

15 T.S. Koritsanszky, P. Coppens, ''Chemical applications of X-ray charge-density analysis''. *Chem. Rev.*, **2001**, *101*, 1583–1627.

16 P. Macchi, A. Sironi, ''Chemical bonding in transition metal carbonyl clusters complementary analysis of theoretical and experimental electron densities''. *Coord. Chem. Rev.*, **2003**, *238–239*, 383–412.

17 W. Scherer, G.S. McGrady, ''Agostic interactions in d^0 metal alkyl complexes''. *Angew. Chem. Int. Ed.*, **2004**, *431*, 782–1806.

18 C. Gatti, ''Chemical bonding in crystals''. *Z. Kristallogr.*, **2005**, *220*, 399–457.

19 C. Lecomte, E. Aubert, V. Legrand, F. Porcher, S. Pillet, B. Guillot, C. Jelsch, ''Charge density research from inorganic and molecular materials to proteins''. *Z. Kristallogr.*, **2005**, *220*, 373–384.

20 R.F.W. Bader, P.M. Beddall, ''Virial field relashionship for molecular charge distributions and spatial partitioning of molecular properties''. *J. Chem. Phys.*, **1972**, *56*, 3320–3329.

21 R.F.W. Bader, H.J.T. Preston, ''The kinetic energy of molecular charge distributions and molecular stability''. *Int. J. Quantum Chem.*, **1969**, *3*, 327–347.

22 R.F.W. Bader, H.J. Essen, ''The characterization of atomic interactions''. *J. Chem. Phys.*, **1984**, *80*, 1943–1960.

23 D. Cramer, E. Kraka, ''A description of the chemical bond in terms of local properties of electron density and energy''. *Croat. Chem. Acta.*, **1984**, *57*, 1259–1281.

24 T.A. Keith, R.F.W. Bader, Y. Aray, ''Structural homeomorphism between the electron density and the virial field''. *Int. J. Quantum Chem.*, **1996**, *57*, 183–198.

25 R.G.A. Bone, R.F.W. Bader, ''Identifying and analyzing intermolecular bonding interactions in van der Waals molecules''. *J. Phys. Chem.*, **1996**, *B100*, 10892–10911.

26 R.W.F. Bader, ''A bond path: an universal indicator of bonded interactions''. *J. Phys. Chem.*, **1998**, *A102*, 7314–7323.

27 A.I. Stash, K. Tanaka, K. Shiozawa, H. Makino, V.G. Tsirelson, ''Atomic interactions in ethylenebis(1-indenyl)zirconium dichloride as derived by experimental electron density analysis''. *Acta Cryst.*, **2005**, *B61*, 418–428.

28 R.W.F. Bader, ''Principle of stationary action and the definition of a proper open system''. *Phys. Rev.*, **1994**, *B49*, 13348–13356.

29 P. Ehrenfest, ''Bemerkungen uber die angenaerte Gueltigkeit der klassischen Mechanik innerhalb der Quantenmechanik''. *Z. Phys.*, **1927**, *45*, 455–460.

30 S. Lundqvist and N.H. March), (Eds.) *Theory of Inhomogeneous Electron Gas*. Plenum Press, New York, **1983**.

31 R.G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, **1989**.

32 R.M. Dreizler, E.K.U. Gross, *Density-functional Theory*. Springer, Berlin. **1990**.

33 D.E. Ellis (Ed.), *Density-functional theory of Molecules, Clusters and Solids*. Kluwer, Dordrecht, **1995**.

34 W. Kohn, A.D. Becke, R.G. Parr, "Density-functional theory of electronic structure". *J. Phys. Chem.*, **1996**, *100*, 12974–12980.

35 M. Springborg (Ed.), *Density-Functional Method in Chemistry and Materials Science*. J. Wiley and Sons, Chichester, **1997**.

36 A. Nagy, "Density-functional theory and application to atoms and molecules". *Phys. Rep.*, **1998**, *298*, 1–79.

37 W. Kohn, "Nobel lecture. Electronic structure of matter–wavefunctions and density functionals". *Rev. Mod. Phys.*, **1999**, *71*, 1253–1266.

38 J.P. Perdew, S. Kurth, "Density functionals for non-relativistic coulomb systems in the new century". *Lecture Notes in Physics*. (Eds. C. Fiolhais, F. Nogueira, M. Marques), Springer, Berlin, Heidelberg, **2003**, 1–55.

39 P. Geerlings, F. De Proft, W. Langenaeker, "Conceptual density-functional theory". *Chem. Rev.*, **2003**, *103*, 1793–1873.

40 U. Von Barth, "Basic density-functional theory – an overview". *Phys. Scr.*, **2004**, *T109*, 9–39.

41 P. Hohenberg, W. Kohn, "Inhomogeneous electron gas". *Phys. Rev.*, **1964**, *136*, B864–B871.

42 W. Kohn, L. Sham, "Self-consistent equations including exchange and correlation effects". *Phys. Rev.*, **1965**, *140*, A1133–A1138.

43 D. Jayatilaka, D.J. Grimwood, "Wavefunctions derived from experiment. I. Motivation and theory". *Acta Cryst.*, **2001**, *A57*, 76–86.

44 N. Hansen, P. Coppens, "Testing aspherical aton refinements on small-molecule data sets". *Acta Cryst.*, **1978**, *A34*, 909–921.

45 F.L. Hirshfeld, "Difference densities by least-squares refinement fumaramic acid". *Acta Cryst.*, **1971**, *B27*, 769–781.

46 E.V. Parini, V.G. Tsirelson, R.P. Ozerov, "The model of multipole anharmonic pseudoatom in precision X-ray structural analysis". *Sov. Phys. Cryst.*, **1985**, *30*, pp. 497–502.

47 A. Volkov, Y. Abramov, P. Coppens, C. Gatti, "On the origin of topological differences between experimental and theoretical crystal charge densities". *Acta Cryst.*, **2000**, *A56*, 332–339.

48 A. Volkov, P. Coppens, "Critical examination of the radial functions in the Hansen–Coppens multipole model through topological analysis of primary and refined theoretical densities". *Acta Cryst.*, **2001**, *A57*, 395–405.

49 R. Flaig, T. Koritsanszky, B. Dittrich, A. Wagner, P. Luger, "Intra- and intermolecular topological properties of amino acids A comparative study of experimental and theoretical results". *J. Am. Chem. Soc.*, **2002**, *124*, 3407–3417.

50 D.A. Kirzhnits, "Quantum corrections to the Thomas–Fermi equation". *Sov. Phys.*, JETP 5, **1957**, 64–72.

51 M. Brack, B.K. Jennings, Y.H. Chu, "On the extended Thomas–Fermi approximation to the kinetic energy density". *Phys. Lett.*, **1976**, *B65*, 1–4.

52 J.A. Alonso, L.A. Girifalco, "Gradient corrections in the energy density functional". *Chem. Phys. Lett.*, **1978**, 190–191.

53 B. Grammaticos, A. Voros, "Semiclassical approximations for nuclear hamiltonians. I. Spin-independent potentials". *Annals of Physics.*, **1979**, *123*, 359–380.

54 W. Yang, "Gradient correction to Thomas–Fermi theory". *Phys. Rev.*, **1986**, *A34*, 4575–4585.

55 W. Yang, R.G. Parr, C. Lee, "Various functionals for the kinetic energy density of an atom or molecule". *Phys. Rev.*, **1986**, *A34*, 4586–4590.

56 Zh.-Zh. Yang, S. Liu, Y.A. Wang, "Uniqueness and asymptotic behavior of the local kinetic energy". *Chem. Phys. Letts.*, **1996**, *258*, 30–36.

57 P. Fuentealba, "Calculation of the atomic kinetic energy from a density functional virial relationship".

J. Phys., B At. Mol. Opt. Phys., **1997**, *30*, 2039–2045.

58 Yu.A. Abramov, "On the possibility of kinetic energy density evaluation from the experimental electron-density distribution". *Acta Cryst.*, **1997**, *A53*, 264–272.

59 P.W. Ayers, P.R. Parr, A. Nagy, "Local kinetic energy and local temperature in the density-functional theory of electronic structure". *Int. J. Quantum Chem.*, **2002**, *90*, 309–326.

60 Y. Tal, R.F.W. Bader, "Studies of the energy density functional approach. I. Kinetic energy". *Int. J. Quantum Chem. Symp.*, **1978**, *12*, 153–168.

61 V. Tsirelson, "The mapping of electronic energy distributions using experimental electron density". *Acta Cryst.*, **2002**, *B58*, 632–639.

62 E. Espinosa, E. Molins, C. Lecomte, "Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities". *Chem. Phys. Lett.*, **1998**, *285*, 170–173.

63 E. Espinosa, E. Molins, "Retrieving interaction potentials from the topology of the electron density distribution. The case of hydrogen bonds". *J. Chem. Phys.*, **2000**, *113*, 5686–5694.

64 V. Tsirelson., Yu. Ivanov, "Mapping the local DFT functions using X-ray diffraction data. Sagamore XIII. Conference on the Charge, Spin and Momentum Densities. Book of Abstracts". *Inst. Experim. Phys.*, **2000**, p. 104.

65 V. Tsirelson, "Combining the density-functional theory and experimental electron density". Conference *on Electron Density Electron Density Measurement, Calculation, and Application. Abstracts.* Wuerzburg Univ., Wuerzburg, **2002**, p.6.

66 A.E. Masunov, S.F. Vyboishchikov, *XVI Intern. Crystallogr. Congress. Collected Abstracts.* Beijing, **1993**, p. 380.

67 R.M.-L. Kitanech, "Joint application of electron and X-ray diffraction to study of properties of the binary crystals". *Ph.D. Thesis. Mendeleev University of Chemical Technology*, Moscow. **1995**.

68 E. Espinosa, I. Alkorta, I. Rozas, J. Elguero, E. Molins, "About the evaluation of the local kinetic, potential and total energy densities in closed-shell interactions". *Chem. Phys. Lett.*, **2001**, *336*, 457–461.

69 P. Mori-Sánchez, A. Martín Pendás, V. Luanã, "A classification of covalent, ionic and metallic solids based on the electron density". *J. Am. Chem. Soc.*, **2002**, *124*, 14721–14723.

70 G. Gervasio, R. Bianchi, D. Marabello, "About the topological classification of the metal–metal bond". *Chem. Phys. Lett.*, **2004**, *387*, 481–484.

71 A. Costales, M.A. Blanco, A. Martín Pendás, P. Mori-Sánchez, V. Luaña, "Universal features of the topological bond properties of the electron density". *J. Phys. Chem.*, **2004**, *A108*, 2794–2801.

72 M.A. Spackman, "Hydrogen bond energetics from topological analysis of experimental electron densities. Recognizing the importance of the promolecule". *Chem. Phys. Lett.*, **1999**, *301*, 425–429.

73 V.E. Zavodnik, A.I. Stash., V.G. Tsirelson, De.R. Vries, D. Feil, "Electron density study urea using TDS-corrected X-ray diffraction data quantitative comparison of experimental and theoretical results". *Acta Cryst.*, **1999**, *B55*, 45–54.

74 H. Birkedal, D. Madsen, R.H. Mathiesen, K. Knudsen, H.-P. Weber, P. Pattison, D. Schwarzenbach, "The charge density of urea from synchrotron diffraction data". *Acta Cryst.*, **2004**, *A60*, 371–381.

75 A. Stash, Unpublished work.

76 A. Granovsky, **2000**, http://classic.chem.msu.su/gran/gamess/index.html

77 M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A.

Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, "General atomic and molecular electronic structure system". *J. Comput. Chem.*, **1993**, *14*, 1347–1363.

78 http://www.chemistry.mcmaster.ca/aimpac

79 V.G. Tsirelson, Yu.A. Abramov, V.E. Zavodnik, A.I. Stash, E.L. Belokoneva, J. Stahn, U. Pietsch, D. Feil, "Critical points in a crystal and procrystal". *Struct. Chem.*, **1998**, *9*, 249–254.

80 R.F.W. Bader, J.A. Platts, "Characterization of an F-center in an alkali halide cluster". *J. Chem. Phys.*, **1997**, *107*, 8545–8553.

81 L.J. Farrugia, P.R. Mallinson, B. Stewart, "Experimental charge density in the transition metal complex $Mn_2(CO)_{10}$ a comparative study". *Acta Cryst.*, **2003**, *B59*, 234–247.

82 E.A. Zhurova, V.G. Tsirelson, "Electron density and energy density view on the atomic interactions in $SrTiO_3$". *Acta Cryst.*, **2002**, *B58*, 567–575.

83 E.A. Zhurova, V.G. Tsirelson, A.I. Stash, M.V. Yakovlev, A.A. Pinkerton, "Electronic energy distributions in energetic materials NTO and the biguanidinium dinitramides". *J. Phys. Chem.*, **2004**, *B108*, 20173–20179.

84 C.H. Hodges, "Quantum corrections to the Thomas–Fermi approximation and the Kirzhnits method". *Can. J. Phys.*, **1973**, *51*, 1428–1437.

85 H. Lee, C. Lee, R.G. Parr, "Cojoint gradient correction to the Hartree–Fock kinetic- and exchange-energy density functionals". *Phys. Rev.*, **1991**, *A44*, 768–771.

86 C.F. Weizsaecker, "Zur Theorie der Kernmassen". *Z. Phys.*, **1935**, *96*, 431–444.

87 T.A. Wesolowski, F. Tran, "Gradient-free and gradient-dependent approximations in the total energy bifunctional for weakly overlapping electron density". *J. Chem. Phys.*, **2003**, *118*, 2072–2080.

88 E.J. Baerends, O.V. Gritsenko, "A quantum chemical view of density-functional theory". *J. Phys. Chem.*, **1997**, *A101*, 5383–5403.

89 A.D. Becke, "Density functional exchange-energy approximation with correct asymptotic behavior". *Phys. Rev.*, **1988**, *A38*, 3098–3100.

90 J.P. Perdew, K. Burke, M. Ernzerhof, "Generelized Gradient Approximation Made Simple". *Phys. Rev. Lett.*, **1996**, *77*, 3865–3868.

91 J.P. Perdew, Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy". *Phys. Rev.*, **1992**, *B45*, 13244–13249.

92 V. Tsirelson, A. Stash, "On functions and quantities derived from the experimental electron density". *Acta Cryst.*, **2004**, *A60*, 418–426.

93 E. Espinosa, M. Souhassou, H. Lachekar, C. Lecomte, "Topological analysis of the electron density in hydrogen bonds". *Acta Cryst.*, **1999**, *B55*, 563–572.

94 P.R.T. Schipper, O.V. Gritsenko, E.J. Baerends, "Kohn–Sham potential and exchange and correlation energy densities from one- and two-electron density matrices for Li_2, N_2, and F_2". *Phys. Rev.*, **1998**, *A57*, 1729–1742.

95 V. Tsirelson, A. Stash, "Determination of the electron localization function from electron density". *Chem. Phys. Lett.*, **2002**, *351*, 142–148.

96 V. Tsirelson, A. Stash, "Analyzing experimental electron density with localized-orbital locato". *Acta Cryst.*, **2002**, *B58*, 780–785.

97 A. Stash, V. Tsirelson, "WinXPRO a program for calculating crystal and molecular properties using multipole parameters of the electron density". *J. Appl. Cryst.*, **2002**, *35*, 371–373.

98 A.I. Stash, V.G. Tsirelson, "Modern opportunities of calculating physical properties of crystals using experimental electron density". *Cryst. Report.*, **2005**, *50*, 209–216.

99 P. Macchi, P. Coppens, "Relativistic analytical wavefunctions and scattering factors for neutral atoms beyond Kr and for all chemically important ions up to I^-". *Acta Cryst.*, **2001**, *A57*, 656–662.

100 S.S. Iyengar, M. Ernzerhof, S.N. Maximoff, G.E. Scuseria, ''Challenge of creating accurate and effective kinetic-energy functionals''. *Phys. Rev.*, **2001**, *A63*, 052508.

101 C. Lee, Zh. Zhou, ''Exchange-energy density functional reparametrization of Becke's formula and derivation of second-order gradient correction.'' *Phys. Rev.*, **1991**, *A44*, 1536–1539.

102 D.P. O'Neill, P.M.W. Gill, ''Benchmark correlation energies for small molecules''. *Mol. Phys.*, **2005**, *103*, 763–766.

11
Topological Analysis of Proteins as Derived from Medium and High-resolution Electron Density: Applications to Electrostatic Properties

Laurence Leherte, Benoît Guillot, Daniel P. Vercauteren, Virginie Pichon-Pesme, Christian Jelsch, Angélique Lagoutte, and Claude Lecomte

11.1
Introduction

The details observable in protein crystal structures depend on experimental diffraction resolution. In X-ray diffraction (XRD), the experimental resolution (Å) is defined by $d = \lambda/(2 \sin \theta)$ where θ is the Bragg angle and λ the wavelength. At the usual resolution for macromolecular crystallography $1.6 < d < 3$ Å, the number of Fourier components $F(h,k,l)$ of the electron density (ED) as obtained from:

$$\rho_{\exp}(x, y, z) = V^{-1} \sum_h \sum_k \sum_l |F(h,k,l)| \exp i\varphi(h,k,l) \exp -[2\pi i(hx + ky + lz)] \tag{1}$$

is not enough for estimation of all the atomic properties. Crystallographic refinement must generally be completed by using a standard stereochemistry for the macromolecule.

Another possibility, instead of using stereochemistry, is to use topological tools. Indeed, topological analysis of the ED can be applied at different resolution levels. For the lower-resolution structures the density maxima often do not correspond to atomic positions but should rather be associated with atomic groups. For example, at ~3-Å resolution, we and others have shown that for protein crystallography there is a correspondence between ED peaks and the backbone or side-chain atoms of the amino acid residues [1–3]. For example, Guo et al. [2] showed that 3-Å resolution ED distributions could be reconstructed from structure factors generated by residue centers. Topological analysis of a medium resolution ED will, therefore, mostly deal with the $(3,-3)$ critical points (CPs), which are the maxima of the ED distribution, and their correspondence with chemical groups to give a geometrically meaningful representation of the protein. Such represen-

tations built on $(3,-3)$ CP properties have, for example, been used in protein–protein and protein–DNA docking applications [4].

At higher resolution, $d < 2.5$ Å, one can observe atomic details and an isotropic atomic Debye Waller factor $B_j = 8\pi^2\langle u_j{}^2\rangle$, where $\langle u_j{}^2\rangle$ is the mean square atomic displacement, can be estimated. The dynamic structure factor can be written:

$$F_{dyn}(\mathbf{H}) = \sum_{j=1}^{N_{at}} f_j(\mathbf{H}) \exp 2\pi i(\mathbf{H}.\mathbf{r_j}) \exp -(B_j \sin^2 \theta/\lambda) \tag{2}$$

where N_{at} stands for the number of atoms in a unit cell, and $f_j(\mathbf{H})$ stands for the atomic scattering factors, i.e. the Fourier transform of the free and neutral atom ED (independent atom model, IAM).

When $d < \sim 1.4$ Å, the isotropic temperature factor may be replaced by an anisotropic factor for non hydrogen atoms:

$$F_{dyn}(\mathbf{H}) = \sum_{j=1}^{N_{at}} f_j(\mathbf{H}) \exp 2\pi i(\mathbf{H}.\mathbf{r_j}) \exp[-2\pi^2(h_i h_k U^{ij}\mathbf{a}_i^*.\mathbf{a}_k^*)] \tag{3}$$

where U^{ij} are the tensor elements of the anisotropic atomic displacement and $\mathbf{a_i}^*$ stands for the ith reciprocal dimension of the crystal unit cell.

The accuracy of the resulting $(3,-3)$ CPs located at the atomic positions is good enough to validate deviations from the standard geometry and evidence the ED of some hydrogen atoms, especially those with a small B factor. At a resolution value estimated to be $d < 0.9$ Å, diffraction data, when accurate enough, contain information at a subatomic scale. Thus, information on valence ED distribution may be obtained when the anisotropic thermal displacement values are small enough, corresponding to an equivalent isotropic B factor lower than about 5 Å^2 [5].

The $(3,-3)$ CPs correspond to nuclei and therefore give an atomic description of the protein structure; hydrogen atoms clearly show up. Deviations from the spherical free atom ED, i.e. deviations from the IAM, appear in the dynamic experimental deformation ED maps as ED peaks located on the chemical bonds:

$$\Delta\rho_{\mathrm{dyn}}(\mathbf{r}) = V^{-1}\sum_{H}(|F_{obs}| - |F_{cal}|) \exp(i\varphi_{cal}) \exp(-2\pi i\mathbf{H}.\mathbf{r}) \tag{4}$$

At subatomic resolution, analysis of the total ED thus enables topological description of the atomic properties: atomic basins and their properties may be defined (charge, volume, dipole moment) as transferable pieces to evaluate chemical bonding and ligand or inhibitor–protein interactions [23].

In this review, we will describe topological approaches based on both medium and high-resolution ED representations and try to link them in a more general

way by using topological results to calculate electrostatic properties of a high-resolution protein model by using the transferability principle.

The chapter will be divided in four parts. The first part will focus on the technical details of the approaches used for topological analysis of high and medium-resolution ED distributions. The second part will summarize results related to the topological analysis approach based on our multipolar ED database fragment description (see below). To test both methods on a real example we will focus, in the last two sections, on results obtained by both approaches applied to the human aldose reductase (hAR) structure [6]. The analysis will focus on a selected subset of the protein active site involved in the binding of the $NADP^+$ cofactor adenine moiety. The third part will be dedicated to results of topological analysis of ED at medium resolution and the fourth part will describe applications to the modeling of protein electrostatic properties.

11.2
Methodology and Technical Details

11.2.1
Ultra-high X-ray Resolution Approach

In contrast with the IAM model, in which all atoms are assumed to be neutral, spherical, and independent, the crystal static valence ED is modeled by a sum of multipolar pseudo atom density $\rho_{\mathrm{val}}^{\mathrm{at}}(\mathbf{r})$ located at atomic centers [7–11], while the atomic core density $\rho_{core}(\mathbf{r})$ remains unchanged:

$$\rho_{\mathrm{val}}^{\mathrm{at}}(\mathbf{r}) = \kappa^3 P_{val}\rho_{\mathrm{val}}(\kappa r) + \sum_{l=0}^{l\,\max} \kappa'^3 R_{nl}(\kappa' r) \sum_{m\mp} P_{lm\mp} y_{lm\pm}(\theta, \varphi) \tag{5}$$

where the term ρ_{val} represents the spherically averaged free atom HF valence density. The second term of the summation describes the nonspherical part, in which the radial functions used are of Slater type: $R_{nl}(r) = Nr^{\mathrm{nl}} \exp(-\xi r)$.

The functions y_{lm} are spherical harmonics in real form; the (θ, φ) coordinates are expressed in an atom-centered local axis system which facilitates application of chemical similarity. The refinable terms are the κ and κ' coefficients, which describe the expansion–contraction of the perturbed valence ED [11], and the population terms P_{val} and P_{lm}.

To apply this formalism to ultra high resolution protein diffraction data we have proposed a multipolar data library [12, 13]: high resolution XRD data have been collected in the *Laboratoire de Cristallographie et de Modélisation des Matériaux Minéraux et Biologiques* (LCM^3B), Nancy, France, for a large group of mono or polypeptides to precisely determine the ED distribution of all natural amino acids (neutral or charged). Multipole refinement of the related structure factors enabled the building of an experimental database of multipolar properties for each

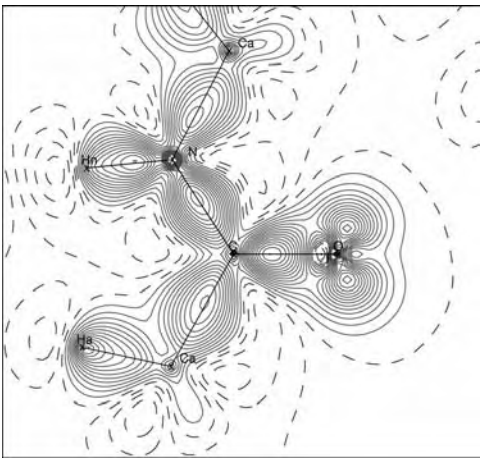


Fig. 11.1 Static deformation ED in the peptide-bond plane obtained with the multipolar ED database. Contour interval 0.05 e Å^{-3}. Full lines: positive values, dashed lines: negative and zero values.

protein-type atom in a given chemical environment [12, 13], from which a specific atom-type nomenclature has been developed. As an example, Fig. 11.1 gives the static electron deformation density:

$$\Delta\rho_{\text{static}}(\mathbf{r}) = \sum_{j=1}^{N_{at}} [\rho_{\text{j mult}}^{\text{at}}(\mathbf{r} - \mathbf{r_j}) - \rho_{\text{j IAM}}^{\text{at}}(\mathbf{r} - \mathbf{r_j})] \tag{6}$$

calculated from the experimental database in the protein main chain HN–C=O peptide plane.

The valence ED distribution of covalent interactions and nonbonding electron pairs is clearly apparent. The database values were shown to be transferable to the protein amino acids and enable calculation of aspherical atomic scattering factors to be used for protein refinement. The validity of these aspherical scattering factors was checked more than 10 years ago [12] and has been confirmed since then by several studies. They have been successfully used to refine ultra-high-resolution protein structures – a scorpion toxin [14], crambin [15], and hAR [16]. As shown in these papers, the use of aspherical scattering factors improves all least-square statistical indices and consequently, leads to more physically meaningful bond distances and thermal anisotropic displacement data. Aspherical features, for example nonbonded density on C=O oxygen atoms, are also taken into account [17]. This finding, which was not surprising because it had already been reported by us and others from porphyrin [18] and naphthalene-type compound [19] charge density studies, led us to the development of the MoPro package of crystallographic programs [10]. In MoPro, for any chemical type of atom belonging to a protein type molecule the corresponding aspherical scattering factor is automatically assigned and these scattering factors can then be further used in the refinement process. Such generalized refinement, which is not

necessarily a *charge density* refinement, because the aspherical scattering factors can be fixed, should be generalized to all small molecule refinements in the future. Furthermore, as shown below, this atom transferability enables estimation of electrostatic properties at least as accurate as when conventional force fields are used [16, 20].

Another approach, based on high-level theoretical calculations, was later proposed by P. Coppens' group [21]. The objective of their theoretical database is different from ours because the also transferable multipolar data will be used to estimate properties such as protein–protein or protein–ligand interaction energies based on theoretical ED reconstruction whereas our approach is based on experimental ultra-high-resolution XRD data. Comparison of our experimental database and Coppens' theoretical database is in progress and will be published soon. A similar project has been recently proposed by Lüger and coworkers [22].

To summarize, we have at our disposal an experimental database which contains data for a set of transferable multipolar pseudo atoms that can be used for protein refinement and estimation of electrostatic properties. These transferable pseudo atoms have their own topology and may therefore be regarded as experimental topological atoms as defined by Bader [23]. In this work, topological analysis of the ED of those fragments has been performed according to the QTAIM theory. The corresponding data, i.e. CP, ρ_b, $\nabla^2\rho_b$ and λ_i calculated using the NewProp software [24], are discussed in Section 11.3.

11.2.2 Medium-resolution Approach

11.2.2.1 Promolecular Electron Density Distribution Calculated from Structure Factors

Several approaches are available for modeling protein ED maps at different levels of resolution. One consists in calculation of structure factors $F(\mathbf{H})$ from the atomic coordinates $\mathbf{r}$ of the system and the IAM atomic scattering factors $f(\mathbf{H})$ and then in the application of a Fourier transformation (FT) algorithm to generate an ED distribution $\rho(\mathbf{r})$ (Eq. 1). This procedure can be completed using any crystallographic refinement program package, for example XTAL [25].

The set of reciprocal space vectors $\mathbf{H}$ that is actually considered in the FT calculation is always finite and determines the crystallographic resolution d of the map.

As already mentioned, in a high-resolution map the density maxima actually correspond to atomic positions whereas in a lower-resolution map they are, instead, associated with groups of atoms. For a protein structure, at a resolution of approximately 3 Å, there is good correspondence between the ED peaks and the backbone or the side-chains of the amino acid residues [1–3]. Guo et al. [2], for example, showed that 3-Å ED distributions could be reconstructed from the so-called "globbic" structure factors associated with residue centers.

Following QTAIM, one can locate $(3, -3)$ CPs, i.e. ED maxima, using the program ORCRIT [26, 27]. In this sense, statistical studies that were conducted on a set of 140 highly idealized protein structures (resolution = 2.85 Å, no hydrogen

atoms, XTAL default overall isotropic thermal variable $\langle u^2 \rangle = 0.035\,\text{Å}^2$, overall atom occupancy = 1) led to macromolecular models built on CP properties that could be used in protein–protein and protein–DNA docking applications [4].

11.2.2.2 Promolecular Electron Density Distribution Calculated from Atoms

Another approach used to model medium-resolution ED distributions consists in smoothing of an analytical representation of the ED. A simple way to model a molecular ED distribution function is to consider that it is the summation over individual atomic contributions ρ_a.

The use of molecular properties such as the ED reconstructed in the framework of a promolecular representation has several essential advantages. First, as for the first method described above computation times are substantially reduced compared with, for instance, ab initio quantum-mechanical calculations. This is especially appealing when studying macromolecular systems [28].

In their works related to the promolecular atom shell approximation (PASA), Amat et al. [29] used atomic Gaussian ED functions that were fitted on 6-311G atomic basis set results (coefficients and exponents can be downloaded from http://iqc.udg.es/cat/similarity/ASA/funcset.html). A molecular or promolecular ED distribution is thus a sum over atomic Gaussian functions wherein expansion coefficients are positive to preserve the statistical meaning of the density function in the fitted structure:

$$\rho_a(\mathbf{r} - \mathbf{R}_a) = Z_a \sum_{i=1}^{3} w_{a,i}[(2\varsigma_{a,i}/\pi)^{3/4} e^{-\varsigma_{a,i}|\mathbf{r}-\mathbf{R}_a|^2}]^2 \tag{7}$$

where $w_{a,i}$ and $\zeta_{a,i}$ are the fitted terms.

In our approach to generate smoothed 3D ED functions, ρ_M is directly expressed as the solution of the diffusion equation according to the formalism presented by [30]:

$$\rho_{a,t}(\mathbf{r} - \mathbf{R}_a) = \sum_{i=1}^{3} s_{a,i} \quad \text{where } s_{a,i} = \alpha_{a,i} e^{-\beta_{a,i}|\mathbf{r}-\mathbf{R}_a|^2} \tag{8}$$

with:

$$\alpha_{a,i} = Z_a w_{a,i} (2\varsigma_{a,i}/\pi)^{3/2} (1 + 8\varsigma_{a,i} t)^{-3/2} \quad \text{and} \quad \beta_{a,i} = \frac{2\varsigma_{a,i}}{(1 + 8\varsigma_{a,i} t)} \tag{9}$$

where t is the ED smoothing degree. It has previously been demonstrated that t is actually equivalent to the overall isotropic temperature factor B, i.e., more precisely [31]:

$$2t = B = 8\pi^2 \langle u^2 \rangle \tag{10}$$

The peaks in a smoothed ED map can obviously be located with ORCRIT, in the same way as for maps generated with the software XTAL. Because the analytical expression of $\rho(\mathbf{r})$ is known, however, peaks can also be located analytically,

without calculation of an ED grid [32]. To follow the pattern of local maxima in a molecular ED distribution, as a function of the degree of smoothing, an algorithm initially described by Leung et al. [33] was implemented. The different steps of that merging/clustering algorithm, described elsewhere [32], consist in following the trajectory of the ED maxima, $\mathbf{r}_{peak}(t)$, in a progressively smoothed ED distribution function:

$$\mathbf{r}_{peak}(t + \Delta t) = \mathbf{r}_{peak}(t) + \frac{\Delta}{\rho_{peak}(t)} \nabla\rho_{peak}(t + \Delta t) \tag{11}$$

where $\nabla\rho_{peak}$ is the density variation and Δ is a predetermined constant value. The trajectory search is stopped when $\nabla\rho_{peak}(t)$ is lower than or equal to a limit value $grad_{lim}$.

In all examples treated so far, the settings were always $\Delta = 2.80 \times 10^{-5}$ Å^2, the number of iterations 2000, and $grad_{lim}$ was set equal to 1.275×10^{-4} e Å^{-4}. The results obtained using that algorithm are the location of the local maxima (peaks), their density, eigenvectors, and Laplacian values, and the atomic content of all fragments, at each value of t between 0 and t_{max} (Fig. 11.2).

11.2.3
A Test System – Human Aldose Reductase

The two topological approaches described above are applied in this section to the hAR structure, solved at a subatomic resolution of 0.66 Å as previously described by Howard et al. [6]. It consists of 316 amino acid residues bonded with the cofactor $NADP^+$ and the inhibitor IDD594. There are 617 water sites and two citrate ions. The whole complex crystallizes in space group $P2_1$ with cell data a = 49.43, b = 66.79, c = 47.40 Å, $\beta = 92.40°$. The full multipolar refinement of hAR is being performed and will be reported elsewhere [34]. In cases of chronic hyperglycemia, aldose reductase is known to reduce part of the excess of glucose to sorbitol [35], accumulation of which in cells leads to long-term diabetes diseases, for example cataracts or nephropathies. hAR inhibition is, therefore, a well researched pharmacological target.

To test the different approaches of ED topological description and to apply these methods in the framework of analysis of electrostatic properties, a small structure subset was selected within the hAR model. It consists of the first amino acid layer surrounding the adenine moiety of the $NADP^+$ cofactor (Fig. 11.3). This subset, referred to below as the "adenine binding site", consists of ten amino acids (160 atoms) of several types:

- four nonpolar residues (LEU212; PRO215; LEU228; ALA245);
- two polar residues (SER263 and ASN272);
- two formally +1 electron (e) positively charged amino acids (LYS262 and ARG268); and
- two formally −1 e negatively charged amino acids (ASP216 and GLU271).

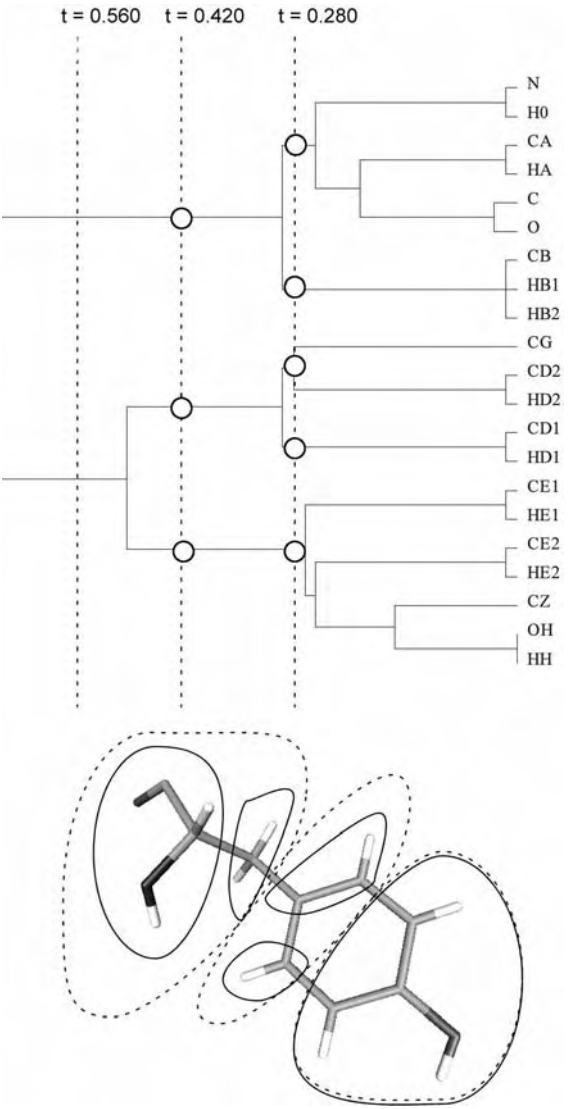


Fig. 11.2 (Top) Dendrogram depicting the results of the hierarchical merging/clustering algorithm applied to the PASA ED distribution of TYR39 of the hAR structure. Results for different values of t (0.280, 0.420, and 0.560 Å^2) are emphasized using vertical lines. The corresponding ED peaks are symbolized using open circles. (Bottom) 2D molecular representation of the information contained in the top figure. Fragments corresponding to ED peaks are represented at $t = 0.280$ (plain lines) and 0.420 (dotted lines) Å^2.

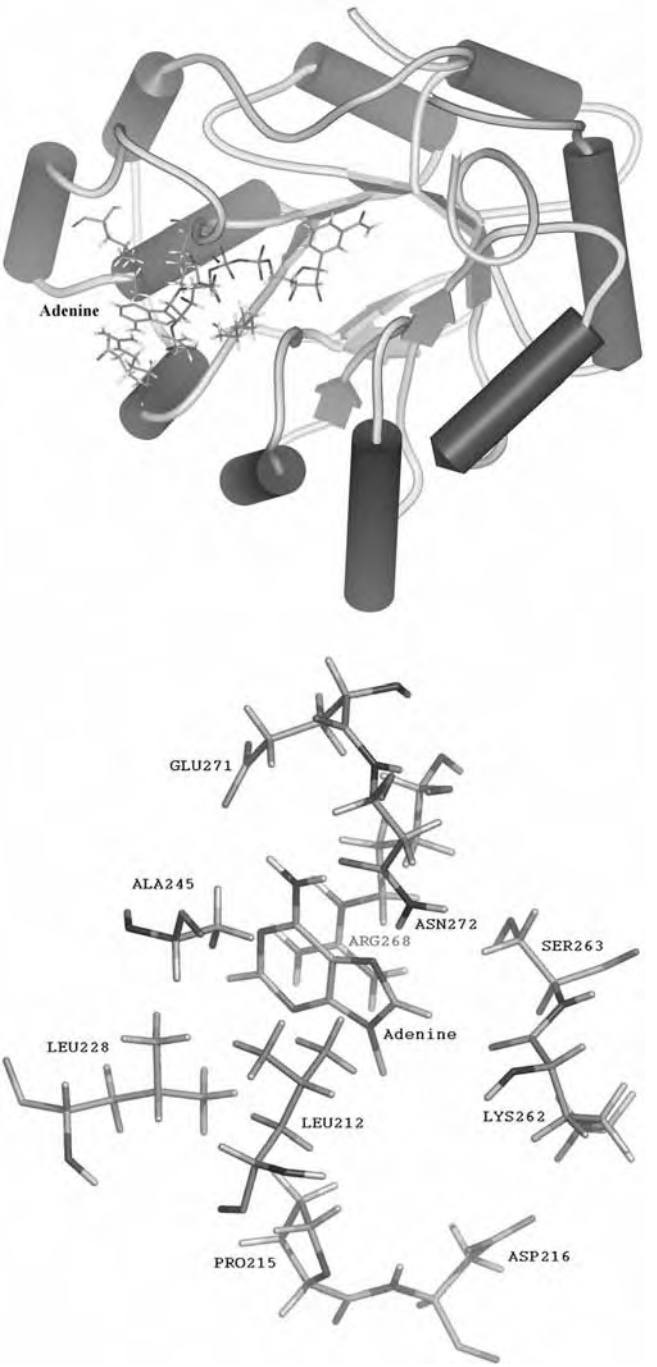


Fig. 11.3 (Top) Schematic view of the hAR structure with the adenine-binding-site amino acids selected for electrostatic potential calculations explicitly represented. (Bottom) Close view of the adenine binding site. It consists of the ten amino acids of the hAR structure directly surrounding the adenine moiety of the $NADP^+$ cofactor.

All electrostatic computations described in the last part of this chapter will take into account this substructure selection only (including hydrogen atoms), thus neglecting the contribution from all other atoms of the protein model. In accordance with the formal charges content of the adenine binding site, global electroneutrality of the fragment will be ensured in all electrostatic potential computations.

11.3
Topological Properties of Multipolar Electron Density Database

As described above, an experimental database of pseudo atom ED is being completed in Nancy [13]. It currently includes all chemical types of atom involved in protein structures and some nucleic acids.

Topological analysis of the ED of the fragments built from our database has been performed as follows using NewProp [24, 36]. The ED of eleven peptides has been reconstructed at the crystallographic experimental geometry using database values to provide some statistical insight of transferability. The QTAIM atom charges Q and volumes V were then calculated for all 101 atom types in the database. When possible, the values were averaged over similar atoms occurring in the molecules stored in the database. For example, for the peptide moiety HN–$H_\alpha C_\alpha$–C=O, C atoms occur 22 times in the database and C_α atoms occur 15 times. Figure 11.4 gives a 2D representation of the atom types (Q versus V). One can see that these two atomic properties can be classified according to their chemical function or/and atomic neighbors. H(X) atoms (where X stands for the bonded heavy atom) can be grouped in three clusters [H(O), H(N)], H(C), and H(S), the largest charge of +0.8 e corresponding to the smaller volume (less than 1 Å^3), depending on the electronegativity of the X atom (or H atom acidity).

The same conclusion can be drawn for the carbon atoms C(O), C(N), or the oxygen O(C), and O(H) atoms with charges as high as +1.25 e and −1.00 e for C and O, respectively.

The C(C) and C(H) atoms (named C in Fig. 11.4) may have a positive or negative charge between ±0.5 e associated with large volume differences (from 6 to 15 Å^3) but no further classification shows up; this is also valid for the N atom but with a less varying charge (−1.1 e to −1.4 e) associated with a large volume (from 10 to 17 Å^3).

The clustering of the topological properties can also be observed by inspection of Table 11.1, which gives the topological characterization of some atom types stored in the database, after transfer to the following moieties:

- the peptide plane HN–$H_\alpha C_\alpha$–C=O, calculated at the experimental geometry of the tyrosine–glycine peptide bond in Leu–Enkephalin [41, 37] compound.
- the aromatic group of the tyrosine amino acid [38].

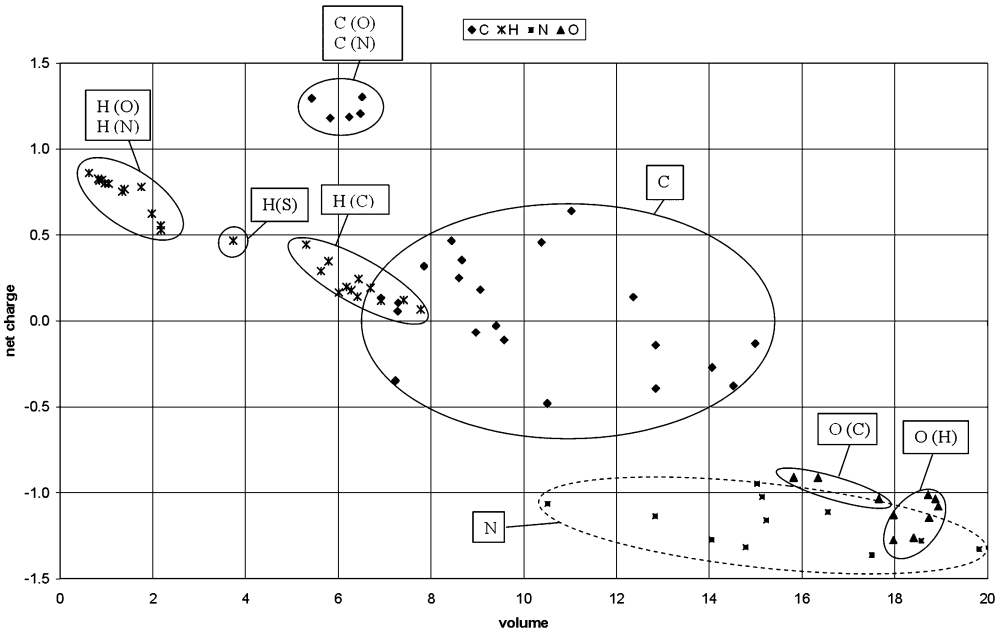


Fig. 11.4 Relationship between net atomic QTAIM charges Q (e) and the atomic basin volume V (Å^3) for all the atom types stored in the multipolar ED database.

The corresponding values theoretically obtained by Matta and Bader [39] are also provided for comparison. The authors used, after geometry optimization, an HF procedure (Gaussian94 software [40] with a 6-311++G** basis set) to calculate the ED of the 24 amino acids in their nonzwitterionic form $H_2N–C_\alpha H_\alpha(R)–COOH$. When available, values described in Coppens' theoretical database [21] for nonhydrogen atoms of the peptide plane moiety are also reported.

The BCP positions in X–H covalent bonds depend on the nature of the X atom, which determines the electron population of the hydrogen atom: hence d_2 (H–BCP distance) changes by 41% when going from H(C) to H(O) (0.34 to 0.20 Å) whereas d_1 (BCP–X distance) changes by 4% (0.74 to 0.77 Å) only; this is also in proportion to the X–H distance. This competition between X and H atoms does not show up when C–N and C–O bond topological properties are compared – even with large electronegativity differences, they both belong to the same cluster but d_1, d_2, or d_1/d_2 do not suggest the same (Q, V) couple.

Comparison of experimental values with Matta's theoretical values reveals good agreement for ρ_b, but with an almost systematic trend – the theoretical ED at BCP is approximately 10% larger than the experimental value. This behavior is less pronounced when experimental database ρ_b values are compared with the Coppens' theoretical values for the peptide plane moiety.

Table 11.1 Topological characterization of the electron density at the BCP.[a]

d_1	d_2	ρ_b	$\nabla^2\rho$	λ_1	λ_2	λ_3	
Peptide group HN–$H_\alpha C_\alpha$–C=O							
0.749	**0.775**	**1.70**	**−12.58**	**−11.90**	**−10.88**	**10.19**	**C_a–C**
0.765	*0.765*	*1.75*	*−17.40*				*C–C (saturated)*
		1.77	−12.9	−12.5	−11.3	10.9	
0.640	**0.818**	**1.89**	**−10.72**	**−13.28**	**−12.80**	**15.36**	**C–N**
0.519	*0.913*	*1.94*	*−20.36*				*C–N*
		1.66	−6.8	−11.40	−10.50	15.1	
0.486	**0.754**	**2.83**	**−28.50**	**−26.19**	**−23.33**	**21.02**	**C=O**
0.397	*0.795*	*2.93*	*2.41*				*C=O*
		2.78	−26.90	−24.10	−22.10	19.3	
0.703	**0.381**	**1.67**	**−14.47**	**−14.79**	**−14.52**	**14.84**	**C_a–H_a**
0.688	*0.396*	*1.98*	*−26.48*				*C–H (saturated)*
0.774	**0.255**	**2.03**	**−25.19**	**−26.58**	**−24.86**	**26.25**	**N–H**
0.745	*0.256*	*2.39*	*−45.11*				*N–H*
Tyrosine aromatic group –CCar– $(CHar–Har)_4$–OHTyr–HOTyr							
0.696	**0.697**	**2.13**	**−19.58**	**−16.08**	**−13.55**	**10.06**	**CHar–CHar**
0.692	*0.694*	*2.18*	*−24.38*				
0.692	**0.700**	**2.12**	**−19.31**	**−15.91**	**−13.48**	**10.08**	**CCar–CHar**
0.686	*0.706*	*2.17*	*−24.19*				
0.694	**0.700**	**2.14**	**−19.91**	**−16.58**	**−13.55**	**10.22**	**CHar–COar**
0.646	*0.741*	*2.20*	*−25.42*				
0.564	**0.796**	**2.17**	**−15.09**	**−16.94**	**−16.47**	**18.33**	**COar–OHTyr**
0.435	*0.920*	*1.94*	*−0.24*				
0.737	**0.339**	**1.77**	**−18.61**	**−16.90**	**−15.78**	**14.06**	**CHar–Har**
0.682	*0.394*	*1.97*	*−26.23*				
0.765	**0.205**	**2.05**	**−25.88**	**−33.37**	**−32.36**	**39.85**	**OHTyr–HOTyr**
0.174	*0.773*	*2.59*	*−68.80*				

[a] ρ_b (e Å^{-3}) and $\nabla^2\rho$ (e Å^{-5}) are the ED and its Laplacian at the BCP; λ_1, λ_2, λ_3 (e Å^{-5}) are the eigenvalues of the Hessian matrix of ρ; d_1 and d_2 (Å) are the distances from the BCP to the first and second atoms defining the bond. The first line corresponds to the multipolar database (**bold**). When available, the second line gives the results of Matta and Bader [39] (*italics*). For the peptide group, when present, the third line gives values from Coppens' theoretical databank [21].

The dependence of (Q, V) values on the nature of the peptide is shown in Table 11.2, which gives the topological properties of the C and O atoms of the peptide C=O group as determined over all peptide molecules stored in the database. Average values and standard deviations are also reported. The C and O atomic multipolar ED data are the same as extracted from the database, irrespective of the type of peptide, and the resulting (Q, V) values only depend on the nature of the side-chain or on its conformation. Because the local C=O geometry is

Table 11.2 Atomic QTAIM charge, Q (e), and basin volume, V (Å^3), obtained from topological analysis of the multipolar ED database. The basin volumes have been defined by inter atomic boundaries based on zero flux surfaces. The results are only given for the carbonyl C=O atoms in several peptide crystals.

Carbonyl Molecule[a]	Carbon			Oxygen		
	Atom	Q	V	Atom	Q	V
actr	C_1	1.118	6.645	O_1	−0.874	16.470
actr	C_2	1.152	6.028	O_2	−1.024	18.147
acdelt	C_1	1.155	6.136	O_1	−1.012	17.668
acdelt	C_2	1.234	5.679	O_2	−1.005	17.950
enk	C_1	1.157	5.647	O_1	−1.023	17.680
enk	C_2	1.202	5.531	O_2	−1.044	17.731
enk	C_3	1.196	5.783	O_3	−1.030	19.317
enk	C_4	1.184	6.261	O_4	−1.047	17.051
trig	C_1a	1.215	5.687	O_1a	−1.055	16.960
trig	C_2a	1.205	5.609	O_2a	−1.025	15.959
trig	C_1b	1.214	5.378	O_1b	−1.030	16.335
trig	C_2b	1.209	5.717	O_2b	−1.034	17.232
ygg	C_1	1.165	5.390	O_1	−1.109	18.443
ygg	C_2	1.200	5.466	O_2	−1.062	16.586
gd	C_1	1.184	6.082	O_1	−1.048	17.392
actyr	C_1	1.083	6.445	O_1	−0.992	18.355
gt	C_1	1.221	5.854	O_1	−1.043	17.376
prohis	C_1	1.173	5.719	O_1	−1.045	16.832
prohis	C_2	1.257	5.037	O_2	−1.072	19.927
prohis	C_3	1.198	5.705	O_3	−1.059	18.919
acgln	C_1	1.104	6.412	O_1	−1.065	17.962
alamet	C_1	1.167	5.917	O_1	−1.040	18.235
Average		1.182	5.824		−1.035	17.660
RMSD[b]		0.041	0.380		0.042	0.963
SEM[b]		0.009	0.083		0.009	0.205

[a] enk, Leu–enkephalin [37, 41]; ygg, Tyr–Gly–Gly [38]; gd, Gly–Asp [38]; actr, *N*-acetyl-L-tryptophan [42]; acdelt, *N*-acetyl-α,β-dehydrophenylalanine methylamide [43]; trig, triglycine [44]; actyr, *N*-acetyl-L-tyrosine ethyl ester monohydrate [45]; gt, glycyl-L-threonine dihydrate [46]; alamet: DL-alanylmethionine [47]; prohis, terbutyl-CO-proline–histidine–NHmethyl [48]; acgln, *N*-acetyl-L-glutamine [49].

[b] RMSD, root mean square deviation for the sample ($N = 23$); SEM, $= \mathrm{RMSD}/\sqrt{(N-1)}$, standard error of the mean.

identical for all carbonyl groups (bond lengths differ by 0.01 Å only), and the interatomic surface of the C atom is limited by N, O, and C_α, the variability of the C atomic basin can originate only from the region above and below the peptide plane, i.e. the side-chains. The O atom is bonded to C only, but its basin is also closed by the intermolecular interactions that occur in the crystal. This could explain the greater variability of V for the O atom than for the C atom, as shown in Table 11.2. In contrast, the QTAIM net charges Q show almost no fluctuations, with $Q(C) = 1.181(9)$ e, and $Q(O) = -1.035(9)$ e. In conclusion, the only significant (but small) change is the oxygen atomic basin with $V = 17.66\,Å^3$, $RMSD = 1\,Å^3$, $SEM = 0.2\,Å^3$, because of intermolecular interactions with different H···O hydrogen bond geometries. The QTAIM charges thus seem to be totally transferable and can be tested as simple point charges in electrostatic property calculations (Section 11.5).

The atomic charges on the peptide group and on the tyrosine aromatic ring for different models of the molecular ED are summarized in Table 11.3. Atomic charges presented here are:

- average atomic QTAIM charges, denoted $QTAIM_{EXP}$, obtained by experimental multipolar database ED integration within atomic basins (line 1);
- atomic QTAIM charges, denoted $QTAIM_{THEO}$, as reported by Matta et al. [39] (line 2); and
- atomic charges, denoted Q_{VAL}, directly computed from average P_{val} values stored in the multipolar ED database by using $Q_{VAL} = N - P_{val}$ (line 3)

Comparison of the atomic charge values shows that Q_{VAL} charges are usually much smaller than $QTAIM_{EXP}$ and $QTAIM_{THEO}$ charges, especially for nonhydrogen atoms, i.e., when atoms are associated with large atomic basins. One also observes that, even though the CP of all tyrosine C–C covalent bonds are similar (Table 11.1), their QTAIM charges largely differ and enable very good discrimination of CHar, CCar, and COar atoms (the atom names arise from the multipolar database nomenclature and indicate aromatic carbon atoms linked to two carbon and one hydrogen, to a carbon, or to an oxygen atom, respectively). The basin volumes also enable differentiation of the three types of atom – the more negative the charge, the larger the volume.

11.4
Analysis of Local Maxima in Experimental and Promolecular Medium-resolution Electron Density Distributions

In this part of the chapter results from peak analysis of medium-resolution ED distributions are presented and discussed. Results obtained by use of the so-called promolecular XTAL model are compared with experimental data at the same resolution, i.e. using the observed F_{obs}. All calculated maps were built according to the hAR crystal structure, including hydrogen and solvent atoms with their

Table 11.3 Atomic net charges Q (e), basin volume V (Å^3) for the atoms in the peptide HN–$H_\alpha C_\alpha$–C=O group and in the tyrosine aromatic cycle. The values were obtained by averaging over the n atoms used to build the database.[a]

Peptide group				Tyrosine aromatic group			
Atom type	***Q***	***V***	***n***	**Atom type**	***Q***	***V***	***n***
C	**1.181(9)**	**5.82(8)**	**22**	**CHar**	**−0.270(6)**	**14.06(23)**	**26**
	1.774(6)	*4.59(2)*			*0.019(6)*	*8.27(8)*	
	0.024(7)				−0.155(4)		
O	**−1.035(9)**	**17.66(21)**	**22**	**Har**	**0.244(4)**	**6.43(15)**	**26**
	−1.35(3)	*19.95(7)*			*−0.007(4)*	*7.27(5)*	
	−0.307(3)				0.170(2)		
N	**−1.272(9)**	**14.05(21)**	**21**	**CCar**	**−0.109(11)**	**9.57(24)**	**5**
	−1.160(4)	*16.64(11)*			*−0.005(10)*	*10.32(6)*	
	−0.312(6)				−0.040(34)		
H	**0.752(23)**	**1.33(14)**	**23**	**COar**	**0.466(4)**	**8.44(16)**	**3**
	0.373(4)	*4.52(5)*			*0.521*	*9.113*	
	0.320(5)				0.053(67)		
C_α	**0.135(8)**	**6.92(4)**	**15**	**OHTyr**	**−1.128(11)**	**17.97(56)**	**3**
	0.577(2)	*6.11(2)*			*−1.273*	*18.034*	
	−0.111(10)				−0.461(28)		
H_α	**0.142(2)**	**6.41(19)**	**16**	**HOTyr**	**0.80(10)**	**0.96(51)**	**3**
	−0.003(3)	*6.86(2)*			*0.624*	*2.904*	
	0.196(5)				0.389(18)		
C_αGly	**−0.028(6)**	**9.40(18)**	**12**				
	0.617	*7.305*					
	0.224(9)						
H_αGly	**0.180(2)**	**6.28(18)**	**24**				
	0.009(0)	*6.58(30)*					
	0.201(2)						

[a] First line (**bold**), QTAIM charges ($QTAIM_{EXP}$) from the multipolar database ED; second line (*italics*), $QTAIM_{THEO}$ charges from Matta and Bader [39]; third line, Q_{VAL} charges (multipolar ED database). The estimated standard deviation of the mean is given in parentheses.

refined occupancies. In addition, two experimental $2F_{obs}-F_{calc}$ maps were considered at 2.85 and 3.5 Å resolution.

11.4.1 Experimental and Promolecular Electron Density Distributions Calculated from Structure Factors

Both experimental and a calculated maps were considered for each of the two crystallographic resolution values 2.85 and 3.5 Å as selected from Becue et al.

[3]. As described in Section 11.2.2, the promolecule maps were built with the software XTAL [25], using the experimental atomic positions and thermal data B, whereas the experimental $2F_{obs}-F_{calc}$ maps were obtained directly by Fourier transformation using multipolar phases and structure factors *moduli*. Three of the four maps under study were characterized by the grid intervals 0.581, 0.566, and 0.571 Å, along the unit cell axes a, b, and c, respectively. The experimental map generated at 2.85 Å was calculated using grid intervals equal to 0.499, 0.510, and 0.489 Å.

The software ORCRIT [26] was then applied to these four maps to locate their ED maxima. To remove CPs originating from ripples in the ED distributions because of series-termination errors, a cut-off value was selected to eliminate most of the unidentified low density peaks. This cut-off mainly affects the number of peaks from bound water molecules, as shown in a study about the use of $2F_{obs}-F_{calc}$ maps [50]. This lower limit value and the ED value of the highest peak found in each map are reported in Table 11.4.

In contrast with the hierarchical merging/clustering algorithm based on an analytical derivation of the ED peaks, there is no generation of fragments associated with the peaks. This therefore forbids their identification on the basis of their atomic content. To assign a chemical identification to each protein peak in a given grid, therefore, a list of reference sites was established. For each amino acid residue, n, two centers-of-mass (c.o.m.) locations were calculated, one for the side-chain and the other for the backbone atoms $(C{=}O)_n{-}(N{-}C_\alpha)_{n+1}$. Other selected reference sites were the solvent and heteroatoms of the complex. The peak was then identified by determining its nearest protein, solvent, or heteroatom site.

Table 11.4 Number of peaks in experimental and promolecular XTAL ED maps of hAR at resolution values of 2.85 and 3.5 Å, and mean distances in Å (in parentheses) between the peaks and their nearest amino acid site (main chain or side-chain c.o.m.), or solvent atom site.

Resolution / **No. of peaks**	**2.85 Å**		**3.5 Å**	
	Experimental	**Promolecular**	**Experimental**	**Promolecular**
Main chain	322 (0.599 ± 0.434)	317 (0.548 ± 0.421)	271 (0.997 ± 0.459)	237 (0.995 ± 0.436)
Side-chain	340 (0.893 ± 0.554)	313 (0.870 ± 0.567)	288 (0.868 ± 0.525)	227 (0.770 ± 0.470)
Ligand	7	8	6	5
$NADP^+$	11	10	9	8
Citrate	4	4	6	4
Water	329 (0.478 ± 0.331)	315 (0.408 ± 0.289)	234 (1.193 ± 0.529)	206 (0.940 ± 0.380)
ED range (e $Å^{-3}$)	0.6–4.95	0.8–5.46	0.6–3.86	0.5–3.07

The results reported in Table 11.4 show that the total number of peaks depends, as expected, on the resolution. The number of side-chain peaks is close to the number of backbone peaks. This is especially true at 2.85 Å resolution, where each amino acid residue leads to a backbone and a side-chain peak, as already explained by Leherte et al. [1], Guo et al. [2], and Becue et al. [3]. The mean distances between the peaks and their nearest protein site is indeed shorter at 2.85 Å, except for the side-chains for which the peaks can be located farther from the c.o.m. in long chains. A statistical analysis of the backbone peaks was carried out as a function of the amino acid residue type. It showed that most of the amino acid backbones are represented by one peak only. More precisely, this concerns 90.5% (266/295) and 92.9% (275/296) of the peaks observed in the experimental and promolecular ED maps generated at 2.85 Å, respectively, and 82.9% (189/228) and 95.1% (215/226) in the corresponding ED maps generated at a resolution value of 3 Å. Backbone groups are more often represented by two peaks in the experimental map.

In the same way as for the backbone, side-chain groups lead to most single ED maxima, and most of the residue side-chains represented by two or more peaks can be regarded as medium or large groups. Short side-chain residues containing no heteroatoms (O or S), or only one, are represented by one peak only, i.e. ALA, CYS, GLY, SER, and THR, at $d = 2.85$ Å. At $d = 2.85$ Å, all TRP side-chains, which contain two fused rings, lead to at least two peaks, a trend that is partly verified for TYR side-chains (one aromatic ring and one hydroxyl group).

Finally, there are more discrepancies between experimental and promolecular XTAL at 3.5 Å resolution, e.g. for ARG, CYS, GLY, HIS, LEU, LYS, and PRO. In conclusion, such promolecular models are less predictive (in terms of topology, for instance) at low resolution; to confirm this, however, additional studies would be required.

Three-dimensional representations of the ED distributions of the adenine binding site are displayed in Fig. 11.5. A detailed analysis of the associated peaks is given in Table 11.5. The size and the atomic content of an amino acid residue affect the number of its peaks in a medium-resolution ED map. Table 11.5 shows that the density values at the peak locations are, in contrast, not clearly dependent on amino acid type. Let us also mention, however, that the residues CYS and MET, when present, are an exception because they contain sulfur atoms and lead to higher-density peaks [1, 3].

11.4.2
Promolecular Electron Density Distributions Calculated from Atoms (PASA Model)

The hierarchical merging/clustering algorithm described in Section 11.2.2.2 does not require any calculation of the ED maps. It is based solely on a knowledge of the analytical expression of the promolecular ED function and its first derivative. The decomposition of the protein structure into fragments was achieved at t values ranging from 0 to 0.70 Å^2, i.e. $B = 0$ to 110.6 Å^2, with a step of 0.014 Å^2.

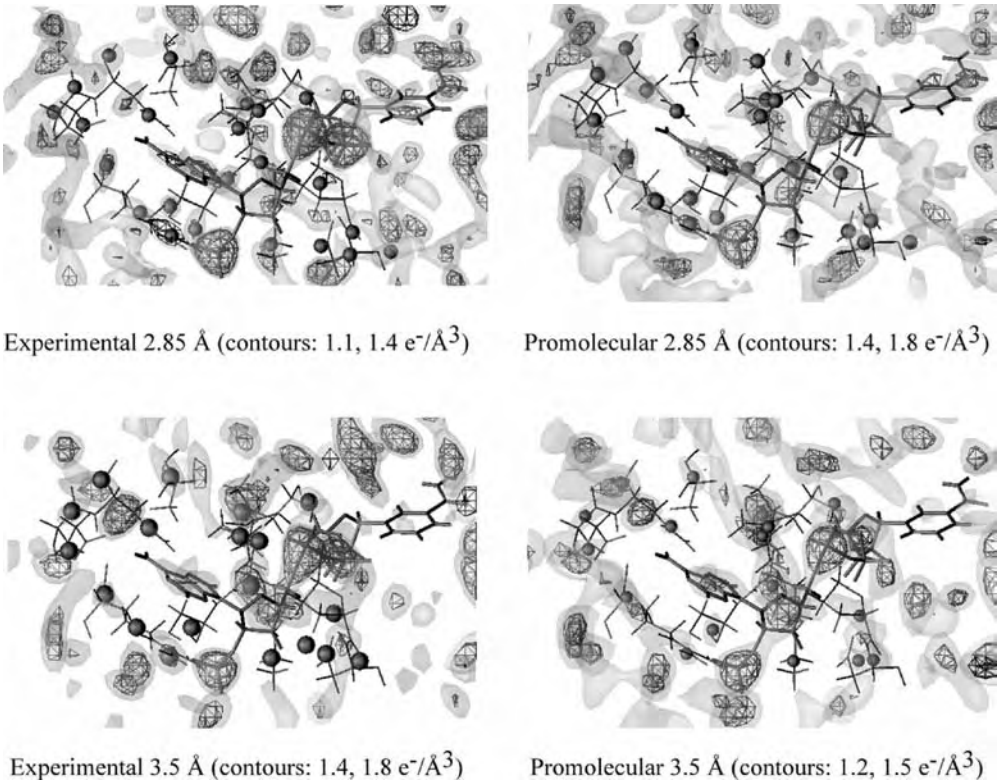

Experimental 2.85 Å (contours: 1.1, 1.4 e⁻/Å3)
Promolecular 2.85 Å (contours: 1.4, 1.8 e⁻/Å3)
Experimental 3.5 Å (contours: 1.4, 1.8 e⁻/Å3)
Promolecular 3.5 Å (contours: 1.2, 1.5 e⁻/Å3)

Fig. 11.5 3D representations of the ten amino acid residues of the adenine binding site of the protein hAR (thin black sticks) and corresponding local maxima (black spheres) observed in experimental and promolecular XTAL ED distributions at resolutions $d = 2.85$ and 3.5 Å. The $NADP^+$ structure is displayed using thick black sticks. Two density iso-contour levels are displayed at a high and a low selected ED value, using triangulated and solid surfaces, respectively.

The evolution of the number of ED maxima as a function of t is shown in Fig. 11.6. This evolution is characterized by several well marked steps associated with the formation of specific fragments. The merging of the atoms to form fragments first occurs between the H atoms and their chemically bonded neighbors at $t = 0.014\,Å^2$. Then, as already shown [31], the C and O atoms of the backbone carbonyl groups begin to merge starting at $t = 0.098\,Å^2$. Between 0.196 and $0.280\,Å^2$, the atoms of the amino acid backbones merge until regular fragment structures, for example (C=O)–N–C_α and (C=O)–N–C_α–C_β (for clarity, H atoms are not shown) are fully created at approximately $t = 0.350\,Å^2$. At this particular value of t the shortest *rms* deviation (0.458 Å) is observed between the coordinates of the backbone peaks and their corresponding c.o.m.

Table 11.5 Identification and density values of the peaks associated with the ten amino acid residues of the adenine binding site of the hAR structure as obtained by CP analysis of experimental and promolecular XTAL ED maps at resolution values of 2.85 and 3.5 Å. ''bak'' and ''sid'' stand for ''backbone'' and ''side-chain'', respectively.

Amino acid site		ρ (e Å^{-3})			
		2.85 Å		3.5 Å	
		Experimental	Promolecular	Experimental	Promolecular
LEU212	bak	1.385	1.737	1.784	1.496
LEU212	sid	1.363	1.781, 0.890	2.103	1.786
PRO215	bak	1.314	1.655		1.405
PRO215	sid	1.779	1.768	1.833	
ASP216	bak	1.452, 1.247	1.823	1.787	1.397
ASP216	sid	1.133	1.618, 1.594	1.548, 1.075	1.376
LEU228	bak	1.409	1.973	1.241	1.476
LEU228	sid	1.199	1.880	1.830	1.711
ALA245	bak	1.841	1.974	2.021	1.428
LYS262	bak	1.681, 1.275	1.972	2.544	1.810
LYS262	sid	1.522	1.754, 1.539	1.551	1.286
SER263	bak	1.599, 1.574	1.831, 1.722	1.789	1.400
ARG268	bak	1.700	2.121	1.745	1.462
ARG268	sid	2.069, 1.524	2.152, 1.750, 1.734	2.099, 1.079	1.415
GLU271	bak	1.255, 1.209	1.547	1.718	1.341
GLU271	sid	0.888	1.546	1.328	1.428
ASN272	bak	1.490	1.864	1.501	1.607
ASN272	sid	1.086	1.726, 1.469	1.688	1.358

The numbers of fragments/peaks obtained at $t = 0.350$ and $0.420\,\text{Å}^2$ are presented in Table 11.6. As already mentioned [31], $t = 0.420\,\text{Å}^2$ enables partitioning of the protein structure into parts located either on the backbone or on the side-chains. At that particular value of t most of the side-chains are represented by one maximum located at an average distance of 0.76 Å from the side-chain c.o.ms (Table 11.6). If the number of backbone peaks is close to the number of amino acid residues, the number of side-chain peaks is relatively smaller than in XTAL maps and the number of solvent peaks is greater. This is because, in XTAL maps, experimental values of B, which differ for each atom and are rather large for the solvent, are considered. In smoothed PASA maps only one value of B (that is equivalent to $2t$) is valid for the whole set of atoms. Let us also mention that the lowest density value in any ED map is equal to 0.0 e Å^{-3}, because of the specific parameterization scheme adopted by Amat et al. [29]. This scheme indeed forces the weight w of the Gaussian functions in the PASA representation (Eqs 7 and 9) to be positive.

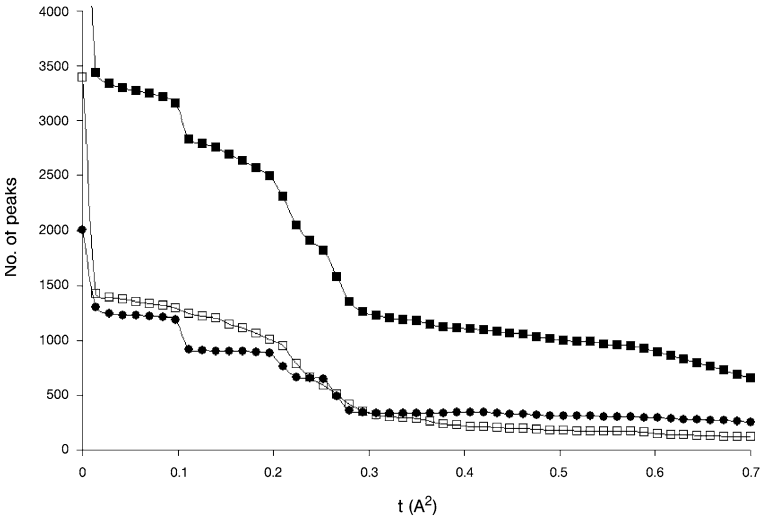


Fig. 11.6 Total number of fragments/peaks (black squares) and numbers of peaks close to any backbone c.o.m. (black spheres) or side-chain c.o.m. (open squares) in the PASA ED distribution of hAR structure as a function of the smoothing term t.

Detailed analysis of the fragment content shows that at $t = 0.350\,\text{Å}^2$, 32% (108/334) of the backbone fragments contain the (C=O)–N–C_α atoms, and 37% (125/334) contain the (C=O)–N–C_α–C_β atoms. With a smoothing value of $0.420\,\text{Å}^2$ these ratios change to 26 (86/331) and 37% (124/331), respectively. Residues are more regularly decomposed into one backbone and one side-chain fragment.

Table 11.6 Number of fragments/peaks in smoothed PASA ED maps of the hAR structure at $t = 0.350$ and $0.420\,\text{Å}^2$, and mean distances in Å (in parentheses) between the peaks and their nearest amino acid site (main-chain or side-chain c.o.m.), or solvent atom site.

Number of peaks	$t = 0.350\,\text{Å}^2$	$t = 0.420\,\text{Å}^2$
Backbone	334; (0.316 ± 0.297)	331; (0.283 ± 0.320)
Side-chains	290; (0.843 ± 0.543)	224; (0.757 ± 0.532)
Ligand	6	6
$NADP^+$	11	9
Citrate	7	6
Water	528; (0.214 ± 0.628)	520; (0.232 ± 0.597)
ED range (e Å^{-3})	0.0–2.895	0.0–2.288

There are two exceptions – the TYR side-chain, which is represented by two fragments, $[C_\gamma–C_\delta]$ and $[(C_\varepsilon)_2C_\zeta–O]$, and VAL, for which backbone and side-chain atoms form only one fragment. The other backbone of the amino acid residues leads either to two fragments, e.g. C=O and $N–C_\alpha–C_\beta$ in for 7.4% (23/311 and 23/308, at 0.350 and 0.420 Å^2, respectively) of the amino acid residues, or to larger fragments containing, for example, C_γ.

As detailed further in this chapter, the atomic fragment content of the ten amino acid residues that constitute the adenine binding site of the protein will be used to determine fragment charges and calculate molecular electrostatic properties. For this particular set of amino acid residues it is apparent the fragment contents are identical, at both $t = 0.350$ and $t = 0.420$ Å^2 (Table 11.7). There are only two exceptions, LYS262 and GLU271, for which two fragments at $t = 0.350$ are merged at $t = 0.420$ Å^2. GLU271 is also a particular example for which the two sites occupied by the side-chain atoms lead to two fragments at $t = 0.350$ Å^2.

For fragments close to a backbone c.o.m., the C=O group belongs to residue n; the other atoms belong to amino acid residue $n + 1$. For example, LEU212_bak is composed of $(C=O)_{212}$ and $(NC_\alpha)_{213}$, and THR244_bak contains $(C=O)_{244}$ and $(NC_\alpha C_\beta)_{245}$. In these two examples, residues 212 and 245 are part of the adenine site whereas 213 and 244 are not.

For a given value of t, all density values are very close to each other. Some fragments are, however, characterized by lower values under two conditions – either there are two fragments for a residue side-chain (LYS262 and GLU271) or the side-chain is composed of C and H atoms only (LEU212 and LEU228).

Comparison of the local density maxima observed in experimental and promolecular XTAL or smoothed ED maps is depicted in Fig. 11.7. It is readily apparent that the strongest similarity (shortest distances) between the two promolecular models and/or the experimental results occurs between the peaks observed in the smoothed PASA maps, at $t = 0.350$ or 0.420 Å^2, and the XTAL/experimental maps at a resolution of 2.85 Å. Indeed, approximately two thirds of the calculated distances lie below 1 Å. For example, 50% and 15% of the distances calculated between the peaks in a map smoothed at $t = 0.350$ Å^2 and the peaks in a promolecular XTAL map at 2.85 Å belong to [0, 0.5 Å[and [0.5, 1.0 Å[, respectively.

11.5
Calculation of Electrostatic Properties from Atomic and Fragment Representations of Human Aldose Reductase

One of the most important applications of this topological analysis is to electrostatic potential properties, as shown, for example, by Muzet et al. [16]. Hence, understanding of electrostatic properties and interaction energies is of prime importance in the prediction of protein–protein, protein–DNA, and drug–receptor recognition and interactions. Systematic calculation of atomic charges by theoretical DFT methods [16] or by experimental electron density analysis [15] can be a

Table 11.7 Atomic content, density values ρ, and charges of the fragments that contain atoms of the hAR adenine binding site as obtained using a hierarchical merging/clustering algorithm applied to the PASA ED distribution. For clarity, bonded H atoms are not shown. "bak" and "sid" stand for "backbone" and "side-chain", respectively.

Nearest c.o.m.		Fragment content	ρ (e Å^{-3})		Fragment charge (e), $t = 0.420$ Å^2		
			$t = 0.350$ Å^2	$t = 0.420$ Å^2	Unit charge	Q_{VAL}	$QTAIM_{EXP}$
PRO211	bak	(C=O)–N–C_α	1.167	1.039		0.059	−0.318
LEU212	bak	(C=O)–N–C_α	1.174	1.039		−0.206	0.111
LEU212	sid	C_β–C_γ–(C_δ)C_δ	0.999	0.891		0.258	0.448
SER214	bak	N–C_α–C_β–C_γ–C_δ	1.181	1.073		0.196	−0.583
PRO215	bak	(C=O)–N–C_α–C_β	1.181	1.046		−0.050	−0.104
ASP216	bak	(C=O)–N–C_α–C_β–C_γ	1.188	1.053		−0.206	0.111
ASP216	sid	C_γ–(O_δ)O_δ	1.181	1.039	−1	−0.833	−0.578
LEU227	bak	(C=O)–N–C_α–C_β	1.188	1.059		0.155	−0.216
LEU228	bak	(C=O)–N–C_α–C_β–C_γ–C_δ–(O_ε)O_ε	1.174	1.046		−0.206	0.111
LEU228	sid	C_γ–(C_δ)C_δ	0.999	0.891		0.162	0.346
THR244	bak	(C=O)–N–C_α–C_β	1.188	1.059		0.208	−0.215
ALA245	bak	(C=O)–N–C_α–C_β	1.181	1.053		−0.206	0.111
PRO261	bak	(C=O)–N–C_α–C_β	1.181	1.039		0.155	−0.216
LYS262	bak	(C=O)–N–C_α	1.174	1.046		−0.050	−0.104
LYS262	sid	C_γ	0.918				
LYS262	sid	C_δ–C_ε–N_ζ	0.985	0.857	1	0.832	0.706
SER263	bak	(C=O)–N–C_α	1.174	1.039		−0.206	0.111
SER263	sid	C_β–O_γ	1.059	0.918		−0.109	−0.347
GLU267	bak	(C=O)–N–C_α–C_β	1.181	1.046		0.215	−0.216
ARG268	bak	(C=O)–N–C_α	1.181	1.053		−0.206	0.111
ARG268	sid	C_γ–C_δ–N_ε–C_ζ–(N)N	1.181	1.039	1	0.853	1.169
ALA270	bak	(C=O)–N–C_α–C_β–$[C\gamma]_{\text{chainB}}$	1.161	1.039		0.135	−0.215
GLU271	bak	(C=O)–N–C_α	1.147	1.033		−0.157	−0.206
GLU271	sid	$[C_\gamma$–C_δ–$O_{\varepsilon1}O_{\varepsilon2}]_{\text{chainA}}$–$[O_{\varepsilon1}]_{\text{chainB}}$	0.992	0.857	−1	−0.737	−0.477
GLU271	sid	$[C_\delta$–$O_{\varepsilon2}]_{\text{chainB}}$	0.817				
ASN272	bak	(C=O)–N–C_α–C_β	1.167	1.039		−0.216	0.111
ASN272	sid	C_β–C_γ–(O_δ)N_δ	1.167	1.019		0.158	0.345

limiting factor when rapid prediction of such interaction patterns is needed. Thus, database retrieval of interaction properties combined with reduced models enable acceleration of such calculations.

In this part of the chapter, adenine binding site high-resolution electrostatic potentials, computed using full multipolar database transferred data and the $QTAIM_{EXP}$ charge model, are compared. These properties are also compared to

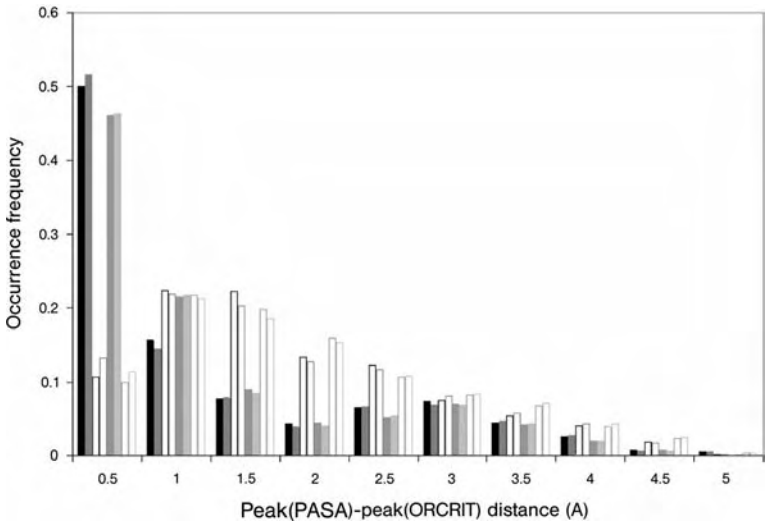


Fig. 11.7 Frequency of occurrence of distances measured between the peaks observed in PASA ED maps of hAR smoothed at $t = 0.350$ (t350) and 0.420 (t420) Å^2, and promolecular XTAL or experimental ED maps at resolution values of 2.85 (R285) and 3.5 (R35) Å. Histogram bars are displayed in the following comparison sequence: t350–R285, t420–R285, t350–R35, t420–R35, t350–R285exp, t420–R285exp, t350–R35exp, t420–R35exp. exp denotes ''experimental''.

the electrostatic potential calculated at medium resolution, using the fragment description of the adenine binding site associated with different point-charge models.

11.5.1
Medium- and High-resolution Approaches of Electrostatic Potential Computations

The high-resolution electrostatic potential generated in the adenine binding site by the protein environment has been computed by VMoPro [10] using, on one hand, the charge-density terms (κ, κ', P_{val}, and P_{lm}, Eq. 5) as transferred from the multipolar database (most accurate description) and, on the other hand, the $QTAIM_{EXP}$ charge model (Section 11.3) [23, 24, 36].

In both cases, to ensure fragment electroneutrality, the total charge of the adenine binding site was set equal to zero by initially correcting atomic charges using the following procedure. The sum of adenine binding site atomic valence populations for neutral and charged atoms are compared. The difference, divided by the number of atoms (*i.e.* 160, including hydrogen atoms), is then subtracted for each atom in the adenine binding site. When using the multipolar database values (which includes Q_{VAL} charges), as already observed by transfer over the full hAR structure [16], the charge increment to be brought to each atom is

very small, +0.008 e, and is thus acceptable as largely below the standard deviations of the database values and below the experimental standard uncertainties for valence populations observed in a charge-density refinement. For $QTAIM_{EXP}$ charges, the amount of charge needed to ensure electroneutrality of the adenine binding site fragment is slightly higher, +0.018 e, but comparable with the multipolar database standard deviations. The resulting high-resolution electrostatic potentials are displayed Fig. 11.8 in two orthogonal planes.

To compute electrostatic potential in the medium-resolution approach, charges attributed to the adenine binding site fragments obtained by analysis of smoothed PASA ED distributions (with a selected smoothing degree of $t = 0.420\,\text{Å}^2$), were calculated by the following procedure.

First, each fragment that contained at least one atom of the adenine binding site was selected. Fragment charges were then calculated as summations over the atomic charges associated with their constituting atoms. Three different sets of atomic charges were considered – the Q_{VAL} charges (Section 11.3, page 298), the $QTAIM_{EXP}$ charges, and unit charges obtained by setting all fragment charges equal to zero, except for the side-chain fragments ARG (+1), GLU (−1), LYS (+1), and ASP (−1). The latter model was used by Becue to calculate an electrostatic scoring index for protein–protein and protein–DNA docking applications [51].

The adenine binding site fragment charges calculated from the three sets of atomic and unit charges described above are given in Table 11.7. This table shows that, except for ARG268, the Q_{VAL} charges of the charged amino acids ASP216, LYS262, and GLU271 are closer to the very simple unit-charge model than those using the $QTAIM_{EXP}$ charge model. For example, ASP216 side-chain fragment net charge is −0.83 e for Q_{VAL} but only −0.58 e for $QTAIM_{EXP}$. The LYS262 Q_{VAL} charge is 0.83 e compared to 0.71 e for the $QTAIM_{EXP}$ charge. For ARG268, the situation is inverted, 1.17 e for $QTAIM_{EXP}$ charges compared with 0.85 e for the Q_{VAL} charges. Accumulation of charges on side-chain fragments is, however, always coherent with the formal charge currently used for such amino acids. For the neutral or polar amino acids present in the adenine binding site model, Q_{VAL} and $QTAIM_{EXP}$ fragment charges are not always consistent, because numerous sign differences occur (Table 11.7). For example, when the backbone fragment obtained by the clustering procedure adopts a −0.21 e negative charge with Q_{VAL} charges, the corresponding $QTAIM_{EXP}$ charge is equal to +0.11 e. Furthermore, the correlation between the signs of Q_{VAL} and $QTAIM_{EXP}$ pseudo atomic charges (without any clustering procedure) is only 63%. $QTAIM_{EXP}$ charges obtained by integration of the experimental multipolar database ED are usually found be more different than the Q_{VAL} charges – positive and negative charges have more pronounced absolute values.

The adenine binding site fragment charges Q_{VAL}, $QTAIM_{EXP}$, and unit charges, which we have located on ED maxima identifying each fragment in the clustering procedure, have been used to generate point-charge calculated electrostatic potentials based on medium-resolution ED topology. They are shown in Fig. 11.9 in the same orientations as defined for the high-resolution electrostatic potentials.

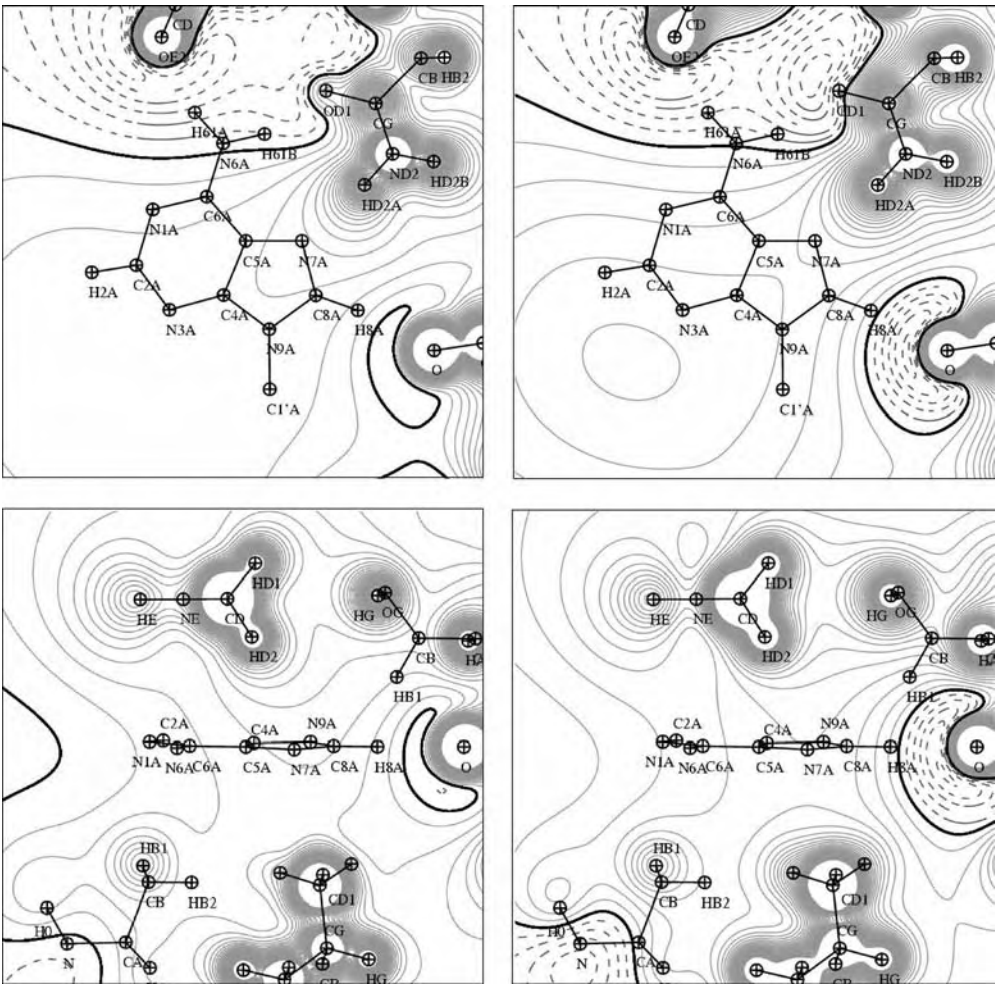


Fig. 11.8 High-resolution electrostatic potentials (e Å^{-1}) calculated using multipolar database values of P_{val} and P_{lm} (left), and $QTAIM_{EXP}$ charge model (right), with the software VMoPro [10]. Contours are displayed in the adenine plane defined by the atoms C5A, N7A, and N6A (top) and in a plane perpendicular to the adenine moiety going through C5A, O_LYS262, and CD_ARG268 atoms (bottom). The interval between contours is 0.05 e Å^{-1}. Negative, positive, and zero contours are displayed in dashed, thin grey, and thick dark lines, respectively.

11.5.2 Electrostatic Potential Comparisons

Initially, despite of the discrepancies pointed out in Section 11.3 between the multipolar database Q_{VAL} charges (from valence population data) and $QTAIM_{EXP}$ charges, comparison of the two high-resolution electrostatic maps calculated with

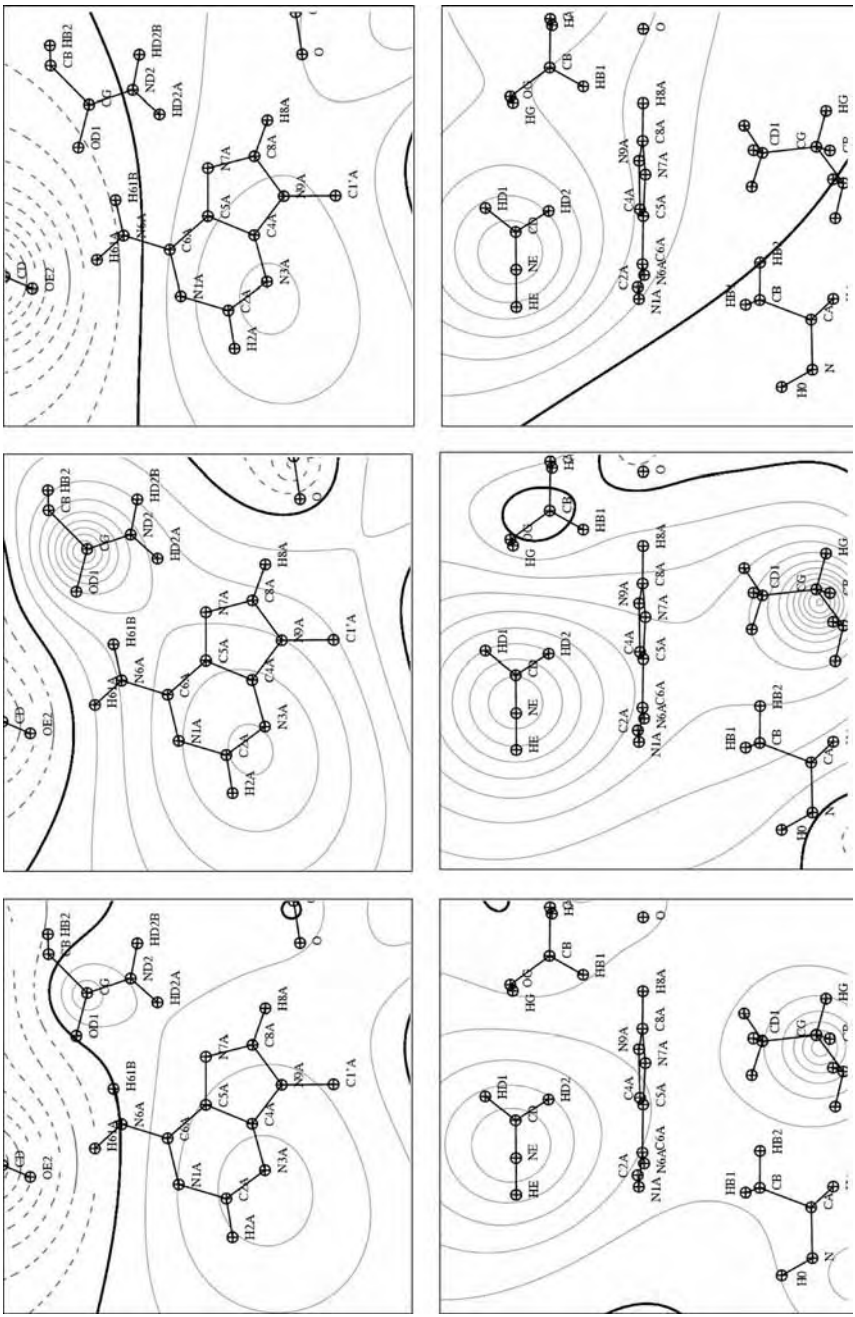


Fig. 11.9 Medium-resolution electrostatic potential calculated from the fragment description of the adenine binding site of hAR using atomic charges from the Q_{VAL} model (left), $QTAIM_{EXP}$ charges model (middle), and unit fragment charges (right). Plane and contour definitions are the same as in Fig. 11.8. Fragments were generated from a PASA ED distribution smoothed at $t = 0.420\,\text{Å}^2$.

VMoPro (Fig. 11.8) reveals surprisingly good qualitative agreement for the adenine moiety binding regions. For example, the electronegative area generated by the group of oxygen atoms OD1_ASN272, OE2, and OE1_GLU271 close to the NH_2 function of the adenine group is very well conserved (Fig. 11.8, top). This is consistent with the electrostatic complementarity expected between the adenine NH_2 group, which has been found to be positively charged in charge-density studies of other adenine-containing compounds [52, 53], and this electronegative region generated by the protein atoms. It is also seen that the adenine double ring is located in a slightly electropositive area in both instances, with atom H8A pointing toward a small region ranging from slightly electronegative (multipolar database values) to strongly electronegative ($QTAIM_{EXP}$ charges), because of the atom O_LYS262. This discrepancy in the area of the H8A adenine atom arises from the more pronounced negative $QTAIM_{EXP}$ charge for the backbone O atom. Very interestingly, however, the region of zero potential (iso-contour 0.0 e $Å^{-1}$) separating electronegative and electropositive regions is always very well conserved, indicating qualitative agreement between the two representations. The main discrepancies between multipolar database and $QTAIM_{EXP}$ descriptions are usually located close to the atoms, because electrostatic potential values computed from data transferred from multipolar database takes directly into account the nonspherical part of the ED in the form of high-order electrostatic moments [54] whose contributions have shorter spatial extent than zero order moments (point charges). Hence, the quantitative difference between the two representations is 0.15 e $Å^{-1}$ close to the O atoms that are hydrogen bonded to the NH_2 adenine group whereas it is only 0.05 e $Å^{-1}$ on its double ring position. When electrostatic potentials are compared in the plane perpendicular to the adenine moiety (Fig. 11.8 bottom), the difference between the electronegative area magnitudes close to the H8A atom is still observed. Another difference arises in the opposite region, where the zero potential iso-contour is shifted toward the protein surface in the $QTAIM_{EXP}$ charges instance, because of their high magnitudes when compared with the Q_{VAL} charges of the multipolar database values.

In the medium-resolution approaches, atomic details in the electrostatic maps are leveled out (Fig. 11.9). The main electrostatic potential features that were pointed out previously are, however, conserved, i.e. the two negative areas close to the $-NH_2$ group and H8A, and a globally positive region at the level of the adenine rings. It is also observed that the position of the zero iso-contour is rather well preserved with the Q_{VAL} charge model whereas it is shifted slightly away from the $-NH_2$ function for the $QTAIM_{EXP}$ charge model when t increases. In maps calculated in the plane perpendicular to the adenine rings, two positive regions appear on each side of the rings. This feature is observed both with the Q_{VAL} and $QTAIM_{EXP}$ charge models but is absent from the results calculated with the unit charge model. One thus concludes that if unit charges are sufficient to globally model the charge distribution of a whole protein structure they are too limited to properly take into account fine details of a receptor site. The Q_{VAL} or $QTAIM_{EXP}$ models should therefore be used.

11.5.3
Electrostatic Interaction Energies

Electrostatic interaction energy values were calculated between the adenine group of $NADP^+$ and the adenine binding site. All three models were considered, at atomic resolution and using a smoothing degree of $t = 0.420\,Å^2$. The atomic charges of the adenine group were those of MacKerell et al. [55]. Interestingly, interaction energy values obtained within the framework of a simple point charge Coulombic interaction depict, in all models, stabilization of the adenine group within the protein cavity. Indeed, electrostatic interaction energies computed in the atomic description using $QTAIM_{EXP}$ charges and multipolar database Q_{VAL} charges are -152.3 and -84.3 kJ mol^{-1}, respectively, whereas the fragment description leads to -74.6, -49.8, and -46.5 kJ mol^{-1} with $QTAIM_{EXP}$, Q_{VAL} and unit charges, respectively. This stabilization is emphasized in the framework of the $QTAIM_{EXP}$ charge model for both atomic and medium-resolution descriptions, because of to their greater magnitudes. Surprisingly, interaction energy obtained for the unit-charge model is very close to that obtained for Q_{VAL} fragment charges, despite discrepancies observed in electrostatic potential representations.

11.6
Conclusions and Perspectives

This chapter presents several concepts and techniques for conducting topological analysis of protein ED distributions at different levels of resolution. It particularly focuses on the transferability of electronic information from high to low-resolution representations.

Several descriptions of ED distributions were considered. First, experimental ED distributions were modeled at a subatomic resolution level ($d < 0.9\,Å$) by use of a multipolar description that enables determination of atomic charges in accordance with the Bader QTAIM approach [24, 36]. Second, a CP analysis technique developed by Johnson [26] was applied to both experimental and promolecular ED distributions at medium resolution ($d \approx 3\,Å$) to locate their ED maxima. In this second approach, promolecular ED grids were obtained by Fourier transformation of calculated structure factors. Third, a smoothing algorithm was applied to a Gaussian promolecular description of the ED distribution to locate its maxima and to define their corresponding molecular fragments by a clustering procedure.

Application of these different approaches were proposed to study the electrostatic properties of the adenine binding site of the hAR structure bound with the cofactor $NADP^+$ [6]. More precisely, the atomic charges that were obtained at a subatomic resolution level were used to calculate medium-resolution fragment charges. Two sets of atomic charges, characterized by overall charge neutrality, were used – database atomic charges directly obtained from the valence popula-

tions and topological QTAIM charges derived from the transferred ED multipolar database.

Topological analysis of ED distributions reconstructed from the multipolar database showed that atomic properties such as charge and volume enabled classification of the atoms according to their chemical function or their chemical environment. For example, atoms bonded to O and N, or to C, formed different clusters. This observation emphasized the effect of the electronegativity of atoms involved in a chemical bond. This clustering of topological properties was described in more detail for the atoms belonging to the peptide group HN–$H_\alpha C_\alpha$–C=O, and to those forming the aromatic group of the amino acid tyrosine. The QTAIM charges seemed to be totally transferable and were thus regarded as simple point charges in electrostatic property calculations.

Molecular electrostatic potentials generated by the adenine binding site in the hAR structure and the electrostatic interaction energies of the adenine moiety with the protein binding site were calculated using several charge models. The two first models were built from the two sets of atomic charges. Each of these two sets was also used to calculate fragment charges. Finally, a fifth charge model was built by assigning formal unit charges to the ARG, LYS (+1) and GLU, ASP (−1) side-chains. Comparison of these five models led to several observations. First, although atomic database and QTAIM charges have similar signs, the QTAIM charges have more pronounced positive and negative values than the database charges. When applied to fragment representations, database and QTAIM fragment charge signs may differ, depending on fragment content. Second, all charge models, atomic or fragment-based, gave similar results, i.e. the adenine moiety is located in a slightly positive area and its $-NH_2$ and H8A group/atom are pointing toward a negative area. It is however shifted away from the adenine group with the topological model when t increases. Fragment representations, up to medium smoothing ($t = 0.420\,Å^2$), are thus acceptable descriptions for modeling electrostatic properties of a protein site. Third, two positive regions occur on either side of the adenine plane, except with the formal unit charge model. In that last instance all four unit charges of the adenine binding site are located on one side only of the adenine rings and do not properly depict its 3D electrostatic properties. Finally, all electrostatic interaction energies adopt the same stabilization effect toward the adenine group within its binding site, even with the formal unit charges. The topologically based electrostatic energy is, however, more negative, because of the more pronounced individual atomic charges.

This work will have a variety of applications. In this chapter, the clustering procedure used to partition a protein structure into fragments is based on progressive smoothing of the protein promolecular ED distribution, i.e. a change in the overall temperature factor. The decomposition of a protein structure as a function of the crystallographic resolution rather than the temperature factor could be considered. This would enable deeper comparison with the CPs obtained using Johnson's approach [26]. For more general application purposes, a database of atomic charges could be combined with a database of protein fragment contents and used to automatically generate reduced protein steric and electrostatic representa-

tions. Such representations are expected to be useful in docking applications, for example, which are nowadays largely used to predict protein–protein, protein–DNA, and protein–ligand interaction patterns.

Acknowledgments

The Université Henri Poincaré – Nancy I, and CNRS are gratefully acknowledged for financial support. A.L. is indebted to the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche for a doctoral fellowship. C.L., B.G., V.P. and C.J. thank Dr A.D. Podjarny for providing us with the hAR crystallographic data. L.L. and D.P.V. thank L. Piela and R. Carbò-Dorca for fruitful discussions. The FNRS-FRFC, the "Loterie Nationale" (convention no. 2.4578.02), and the Facultés Universitaires Notre-Dame de la Paix (FUNDP), are gratefully acknowledged for the use of the Interuniversity Scientific Computing Facility (ISCF) Center.

References

1 Leherte L.; Fortier S.; Glasgow J. and Allen F. H. *Acta Cryst. D* 50 (1994), 155–166.
2 Guo D.-Y.; Blessing R. H.; Langs D. A. and Smith G. D. *Acta Cryst. D* 55 (1999), 230–237.
3 Becue A.; Meurice N.; Leherte L. and Vercauteren D. P. *Acta Cryst. D* 59 (2003), 2150–2162.
4 Becue A.; Meurice N.; Leherte L. and Vercauteren D. P. *J. Comp. Chem.* 25 (2004), 1117–1126.
5 Jelsch C.; Pichon-Pesme V.; Lecomte C. and Aubry A. *Acta Cryst.* D 54 (1998), 1306–1318.
6 Howard E. I.; Sanishvili R.; Cachau R. E.; Mitschler A. et al.; *Proteins Struct. Funct. Bioinformat.* 55 (2004), 792–804.
7 Stewart R. F. *J. Chem. Phys.* 51 (1969), 4569–4577.
8 Hansen N. K. and Coppens P. *Acta Cryst. A* 34 (1978), 909–921.
9 Guillot B.; Viry L.; Guillot R.; Lecomte C. et al.; *J. Appl. Cryst.* 34 (2001), 214–223.
10 Jelsch C.; Guillot B.; Lagoutte A. and Lecomte C. *J. Appl. Cryst.* 38 (2005), 38–54.
11 Coppens P.; Guru Row T. N.; Leung P.; Stevens E. D. et al.; *Acta Cryst. A* 35 (1979), 63–72.
12 Pichon-Pesme V.; Lecomte C. and Lachekar H. *J. Phys. Chem.* 99 (1995), 6242–6250.
13 Pichon-Pesme V.; Jelsch C.; Guillot B. and Lecomte C. *Acta Cryst. A* 60 (2004), 204–208, and references cited therein.
14 Housset D.; Benabicha F.; Pichon-Pesme V.; Jelsch C. et al.; *Acta Cryst. D* 56 (2000), 151–160.
15 Jelsch C.; Teeter M. M.; Lamzin V.; Pichon-Pesme V. et al.; *Proc. Natl. Acad. Sci.* (USA) 97 (2000), 3171–3176.
16 Muzet N.; Guillot B.; Jelsch C.; Howard E. et al.; *Proc. Natl. Acad. Sci.* (USA) 100 (2003), 8742–8747.
17 Guillot B. (2002) *Ph. D. Thesis*, LCM^{3}B Université Henri Poincaré, France.
18 Benabicha F. (1986) *Ph. D. Thesis*, LCM^{3}B Université Henri Poincaré, France.
19 Brock C. P.; Dunitz J. D. and Hirshfeld F. L. *Acta Cryst. B* 47 (1991), 789–797.

20 Pearlman D. A.; Case D. A.; Caldwell J. W.; Ross W. S. et al.; *Comp. Phys. Commun.* 91 (1995), 1–41.

21 Koritsanszky T.; Volkov A. and Coppens P. *Acta Cryst. A* 58 (2002), 464–472.

22 Dittrich B.; Koritsanszky T. and Lüger P. *Angew. Chem. Int. Ed.* 43 (2004), 2718–2721, and references cited therein.

23 Bader R. F. W. (1990) *Atoms in Molecules: A Quantum Theory.* Oxford University Press, Oxford, UK.

24 Souhassou M. and Blessing R. H. *J. Appl. Cryst.* 32 (1999), 210–217.

25 Hall S.; du Boulay D. and Olthof-Hazekamp R. (2002) *The Gnu XTAL System of Crystallographic Software, v. 3.7.2*, http://xtal.sourceforge.net/.

26 Johnson C. K. (1977) *ORCRIT: The Oak Ridge Critical Point Network Program.* Oak Ridge TN, USA.

27 Leherte L. and Vercauteren D. P. *J. Mol. Model.* 3 (1997), 156–171.

28 Gironés X.; Amat L. and Carbó-Dorca R. *J. Chem. Inf. Comput. Sci.* 42 (2002), 847–852.

29 Amat L. and Carbó-Dorca R. *J. Comput. Chem.* 18 (1997), 2023–2039.

30 Kostrowicki J.; Piela L.; Cherayil B. J. and Scheraga H. A. *J. Phys. Chem.* 95 (1991), 4113–4119.

31 Leherte L. *Acta Cryst. D* 60 (2004), 1254–1265.

32 Leherte L.; Dury L. and Vercauteren D. P. *J. Phys. Chem. A* 107 (2003), 9875–9886.

33 Leung Y.; Zhang J.-S. and Xu Z.-B. *IEEE T. Pattern Anal.* 22 (2000) 1396–1410.

34 Guillot B.; Jelsch C.; Podjarny A.; Lecomte C. *Acta Cryst. D*. To be published.

35 Yabe-Nishimura C. *Pharmacol. Rev.* 50 (1998), 21–34.

36 Souhassou M. (1997) *NEWPROP: Computer program to calculate the topological properties of electron density.* Internal Report, LCM^{3}B, Université Henri Poincaré, France, http://www.lcm3b.uhp-nancy.fr.

37 Pichon-Pesme V.; Lecomte C.; Wiest R. and Bénard M. *J. Am. Chem. Soc.* 114 (1992), 2713–2715.

38 Pichon-Pesme V.; Lachekar H.; Souhassou M. and Lecomte C. *Acta Cryst. B* 56 (2000), 728–737.

39 Matta C. F. and Bader R. F. *Proteins Struct. Funct. Genet.* 52 (2003), 360–399.

40 Frisch M. J.; Trucks G. W.; Schlegel H. B.; Gill P. M. W. et al.; (1995) *Gaussian94, Revision B.3.* Gaussian Inc., Pittsburgh PA, USA.

41 Wiest R.; Pichon-Pesme V.; Bénard M. and Lecomte C. *J. Phys. Chem.* 98 (1994), 1351–1362.

42 Souhassou M.; Lecomte C.; Blessing R. H.; Aubry A. et al.; *Acta Cryst. B* 47 (1991), 253–266.

43 Souhassou M.; Lecomte C.; Ghermani N.-E.; Rohmer M.-M. et al.; *J. Am. Chem. Soc.* 114 (1992), 2371–2382.

44 Pichon-Pesme V. and Lecomte C. *Acta Cryst. B* 54 (1998), 485–493.

45 Dahaoui S.; Jelsch C.; Howard J. A. K. and Lecomte C. *Acta Cryst. B* 55 (1999), 226–230.

46 Benabicha F.; Pichon-Pesme V.; Jelsch C.; Lecomte C. et al.; *Acta Cryst. B* 56 (2000), 155–165.

47 Guillot R.; Muzet N.; Dahaoui S.; Lecomte C. et al.; *Acta Cryst. B* 57 (2001), 567–578.

48 Bouhmaida N. (1993) *Ph. D. Thesis*, LCM^{3}B Université Henri Poincaré, France.

49 Pichon-Pesme V. (2001) Personal communication.

50 Minichino A.; Habash J.; Raftery J. and Heliwell J. R. *Acta Cryst. D* 59 (2003), 843–849.

51 Becue A. (2004) *Ph. D. Thesis*, Facultés Universitaires Notre-Dame de la Paix (FUNDP) Namur, Belgium.

52 Guillot B.; Muzet N.; Artacho E.; Lecomte C. et al.; *J. Phys. Chem. B* 107 (2003), 9109–9121.

53 Cunane L. M. and Taylor M. R. *Acta Cryst. D* 53 (1997), 765–776.

54 Ghermani N.; Lecomte C. and Bouhmaida N. *Z. Naturforsch. A* 48 (1993), 91–98.

55 MacKerell Jr. A. D.; Wiorkiewicz-Kuczera J. Jr. and Karplus M. *J. Am. Chem. Soc.* 117 (1995), 11946–11975.

12
Fragment Transferability Studied Theoretically and Experimentally with QTAIM – Implications for Electron Density and Invariom Modeling

Peter Luger and Birger Dittrich

12.1
Introduction

It is interesting to note that as early as 1915, three years after the discovery of X-ray diffraction, Peter Debye made a note that the distribution of electron density should be obtainable from this new experimental method [1]. It took ninety years, however, until, in the autumn of 2005, Philip Coppens' highlight article (entitled "Charge Densities Come of Age") appeared in "Angewandte Chemie" [2]. He pointed out it was a long way from Debye's vision in 1915 to the current stage of electron-density work, that progress was slow, and that the major advances had all occurred in the last decade. The substantial progress made in electron density research in the last few years rests on three pillars:

1. advances in experimental techniques – high-brilliance X-ray sources (synchrotron beamlines), very low temperatures ($T \rightarrow 10$ K), and CCD area detection;
2. theoretical developments, e.g. Bader's QTAIM theory [3], enabling the derivation of quantitative topological data; and
3. computing – the development and distribution of specific computer program systems for all aspects of electronic density work, including refinement, analysis, and visualization of results.

As a result of these simultaneous developments, electron-density studies of entire classes of chemically related compounds or of larger molecules became feasible in a reasonable time. Routine application of electron-density work is in sight.

In this chapter we would like to start with a brief summary of recent experimental advances which enable verification of the transferability of atomic fragments from significant experimental data, a key feature of Bader's QTAIM theory. Understanding of what constitutes chemical similarity led to the introduction of invariom, pseudoatom fragments of electron density that are invariant in transfer from one molecule to another, and to their recent applications. Invariom en-

able replacement of the independent atom model (IAM), and their introduction will possibly add another pillar to high-resolution crystallographic work.

12.2
Experimental Electron-density Studies

12.2.1
Experimental Requirements

The electron density of a chemical structure consists of a large spherical contribution and a very small nonspherical contribution located mainly in the regions of the covalent bonds and some nonbonding (lone pair) regions. Because chemistry happens in the latter regions, one is more interested in the aspherical part of the electron density and observation of these small effects requires very precise experiments. Atoms in crystals are not at rest, being described by well-known displacement terms which decrease as the temperature is reduced. Because thermal smearing should be as small as possible, data collection should be performed at the lowest temperature achievable. At low temperatures, moreover, high-order reflections are more likely to have significant intensities above the background. High-order data are needed not only to improve accuracy and resolution but also to provide sufficient data for the refinement of the increased number of variables of the Hansen and Coppens multipole model [4]. To summarize, several experimental requirements must be met:

- excellent single-crystal quality
- precise intensity-data collection
- high resolution ($\sin\theta/\lambda > 1.0\,\text{Å}^{-1}$ or $d < 0.5\,\text{Å}$)
- hard X-radiation, e.g. MoKα or shorter, synchrotron radiation
- high completeness and redundancy in reciprocal space, hence use of an area detector is preferable
- low temperature ($T \leq 100$ K)

Whether or not a data set is suitable for electron-density determination must be carefully examined. Several criteria can be used:

- conventional figures of merit, R_{int}, Rσ, R(F) or R(F^2), R_w(F), GoF
- the Hirshfeld test [5] – a covalently bonded pair of atoms of comparable mass can be regarded as rigid with respect to thermal motion, so that the displacement parameter components of the atoms in question should be equal in bond direction
- the quality of the residual electron density obtained by Fourier transformation of the difference between observed and multipole-model structure factors – the residual density should be unstructured and low

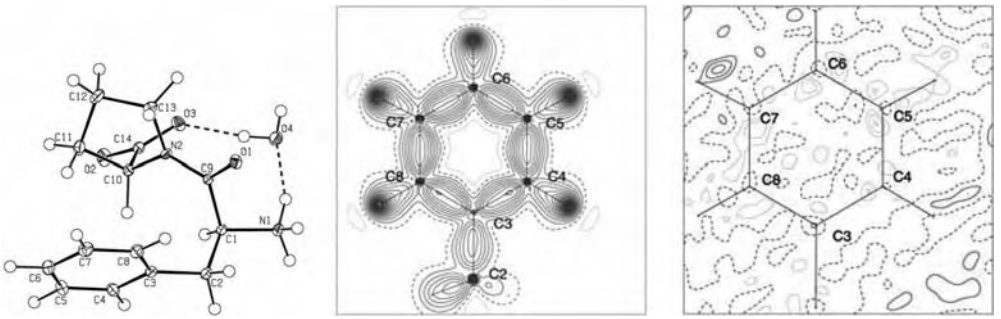


Fig. 12.1 Electron density of the dipeptide L-phenylalanyl–L-proline.H_2O, an example of a proper data set. Tetragonal, $P4_32_12$, $a, b = 8.197$, $c = 41.226$ Å, $V = 2769$ Å^3, $T = 90$ K, MoKα, $(\sin\theta/\lambda)max = 1.20$ Å^{-1}, Smart 1K CCD diffractometer, 112.138 measured reflections, 11.139 symmetry-independent, reflections with Fo $>$ 2.5 σ(Fo):10.120 (91%!), completeness: 99.2%, redundancy: 9.92, R*int* = 2.8%. Multipole model: 653 parameter, R(F) = 2.7%, Rw(F) = 2.3%, GoF = 1.82. Left: Molecular structure with hydrogen bonds to the water molecule; middle/right: static/residual map in the plane of the phenyl ring, contour intervals 0.1/0.05 e Å^{-3}.

Figure 12.1 illustrates an example of a proper electron-density data set measured for the dipeptide L-phenylalanyl–L-proline.H_2O. All criteria mentioned above have been fulfilled. In particular, the residual density is featureless except for a small signal next to one of the hydrogen atoms. The Hirshfeld test was also satisfied in that the maximum difference of displacement parameter components on a bond was ≤ 0.0005 Å^2.

12.2.2 Recent Experimental Advances

For several decades single-crystal X-ray data collection was performed with four-circle diffractometers with point detectors operated in serial detection mode, so that usually no more than 500–1000 reflections could be collected per day. The appearance of CCD area-detection diffractometers in the mid nineties was a major breakthrough. It was shown that this type of detection led to data sets of sufficient accuracy for electron-density work and that the data collection time could be reduced from weeks or months to one or a few days [6–8].

12.2.2.1 Synchrotron Radiation Compared with Laboratory Sources

Because optimum experimental conditions depend not only on detector type and quality but also on the choice of primary radiation and the data-collection temperature, we will discuss these aspects in some detail. On a laboratory scale most crystallographers have a sealed 2.4-kW Mo tube or sometimes a rotating anode available; synchrotron radiation should be considered as a favorable alternative. Synchrotron beamlines provide a bright source of X-radiation with outstanding

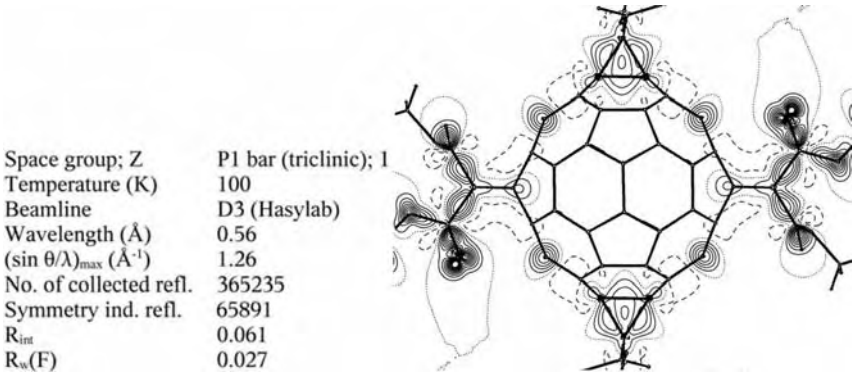


Fig. 12.2 High-resolution synchrotron data set for the fullerene derivative dodecakisethoxycarbonyl-C_{60}-fullerene. Left: summary of the crystallographic data. Right: static deformation density map in an equatorial plane of the C_{60} sphere [9] Copyright 2002 and Reproduction with Permission from American Chemistry Society.

properties. Most important for electron density experiments are the very high intensity and the tunable wavelength, enabling choice of $\lambda \approx 0.5$ Å or even shorter. This enables high-resolution data sets to be collected, if needed also for smaller crystals, for which absorption and extinction problems are minimized.

An example in which synchrotron radiation was essential for collection of a high-resolution data set is illustrated by Fig. 12.2 for a highly substituted C_{60} fullerene derivative.

Fullerenes are usually very unsuitable for experimental electron-density studies because of poor crystal quality and the high mobility of the molecules in the crystal lattice. This is why, for example, there is no electron-density study of an unsubstituted fullerene. After several unsuccessful attempts we were able to grow suitable crystals of the highly substituted T_h-symmetrical derivative dodecakisethoxycarbonyl-C_{60}-fullerene, $C_{102}H_{60}O_{24}$ [10] (Fig. 12.2), which co-crystallized with 1,2-difluorobenzene. On the basis of a data set of more than 350,000 reflections measured in five days at 100 K up to a resolution of $\sin\theta/\lambda = 1.26\,\text{Å}^{-1}$ on beamline D3 of Hasylab/DESY, a properly resolved electron-density distribution and related bond-topological and atomic properties [9, 11] were derived.

If well-diffracting crystals of sufficient size can be grown, high-resolution data collection can be conducted with laboratory equipment. An example is a data set for vitamin B_{12} (Fig. 12.3). Although B_{12} crystals of desired size can be grown rather easily in a variety of solvents, they are extremely unstable when taken out of the mother liquor, being destroyed within seconds. This made careful, but also time-consuming crystal preparation necessary. To avoid the risk of wasting synchrotron beamtime, data were collected by use of a Bruker Smart 1K diffractometer with conventional MoKα radiation from a sealed 2.4-kW tube and were very suitable for further electron-density evaluation. Figure 12.3 shows the static defor-

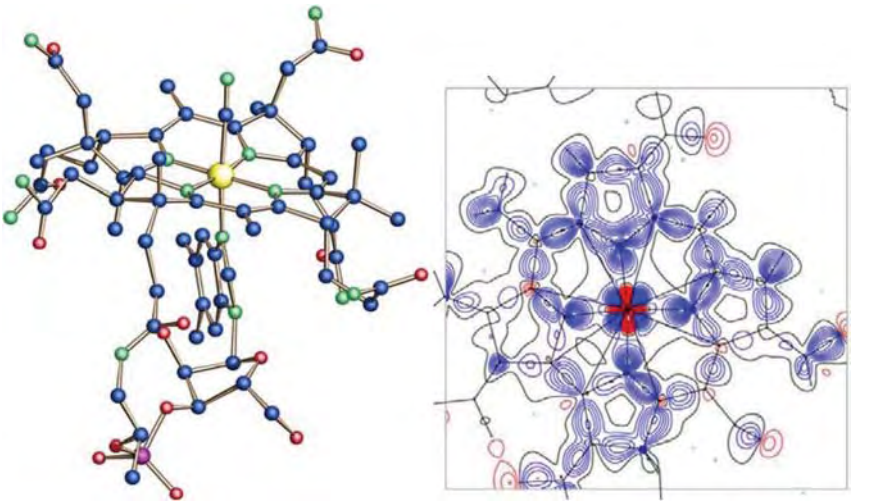


Fig. 12.3 Left: molecular structure of vitamin B_{12} in the crystalline modification cocrystallizing with 12 water and 3 propanol molecules. Summary of crystallographic data [12]: orthorhombic, $P2_12_12_1$, cell volume 8952(3) Å^3, Z = 4, MoKα, T = 100 K, a total of 660,822 reflections collected, $(\sin\theta/\lambda)max = 1.22$ Å^{-1}, 98,204 unique, 78,606 reflections with $I > 2\sigma(I)$ (80%), R*int* = 0.038, R(mul) = 0.03. Right: experimental static deformation density in the plane of the corrin ring. Contour levels 0.1 e Å^{-3}.

mation density distribution in the corrin plane; further derivation of topological properties is still in progress.

12.2.2.2 **Data Collection at Ultra-low Temperatures (10–20 K)**

Low-temperature diffraction experiments are normally conducted at temperatures of approximately 100 K, making use of conventional nitrogen gas-stream devices. It has been shown [13], however, that substantial and, hence, favorable reduction of thermal motion occurs when the temperature is further reduced to 10–20 K. Because of reduction of the displacement parameters, the number of high-order reflections with intensities significantly above the background increases, and this is of utmost importance for electron-density work. These ultra-low temperatures can be achieved either by replacing nitrogen with helium in open-flow gas-stream devices or by using closed-cycle helium cryostats. Open-flow helium systems have the disadvantages of high cost of helium consumption and are, for this reason, rarely used in normal university laboratories.

Closed cycle helium cryostats which, in the double-stage version, reach a temperature of 10–20 K, are a low-running-cost alternative, because no gas is consumed. The cold head of the cryostat, which carries the crystalline sample, must, however, be enclosed in a cylindrical vacuum chamber. When point detectors were used, beryllium or carbon chambers were acceptable, but these result in an intolerable background pattern on CCD area detector frames. This technical

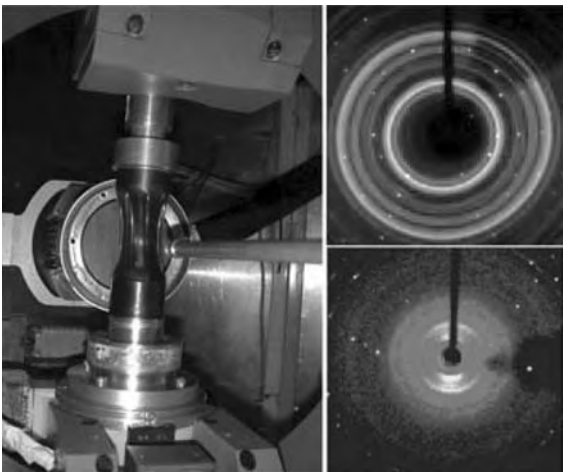


Fig. 12.4 Left: Huber four-circle diffractometer at beamline D3 (Hasylab) equipped with helium cryostat and Bruker Smart 1K CCD area detector. The vacuum chamber wall at the cold head of the cryostat is made of Kapton film. Right: background resulting from different vacuum cylinder materials – beryllium (above) and Kapton film (below) [14] Copyright 2003 and Reproduction with Permission from International Union of Crystallography (IUCr).

problem of the vacuum chamber was recently solved by replacing beryllium with a Kapton film as cylinder-wall material of the vacuum chamber (Fig. 12.4) [14].

This results in a much lower and practically unstructured background and makes the combination of a closed-cycle cryostat with a modern area detector feasible. In a comparative study of four high-order data sets for strychnine single crystals at 100, 25, and 15 K, it was shown that I/σ ratios of high-order reflections improved most favorably when synchrotron radiation and a temperature of 15 K were used [14].

On the basis of current technical developments, experimental advances in high-speed evaluation of electron densities can be expected – data collection periods of hours or even minutes are within reach. This was demonstrated by a 12-h sequence of diffraction experiments recently conducted at the protein crystallography beamline X10 SA of the Swiss light source (SLS) at the Paul Scherrer Institute (Villigen, Switzerland) [15]. It resulted in a total of 400,000 reflections of four high-resolution X-ray data sets for electron density determination. Because of the brilliant primary beam properties, intensities could be observed even for tiny crystals and in very high regions of reciprocal space.

One particular example of this experimental sequence is illustrated in Fig. 12.5, which shows the static map in the purine plane of adenosine after aspherical atom refinement. It was generated from a quick 1 h test dataset of more than 22,000 reflections. The covalent bonding features are properly resolved, even though this data set was acquired in one of the shortest measurements ever conducted in experimental electron-density work.

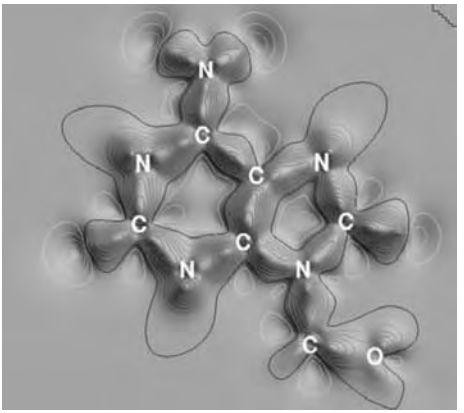


Fig. 12.5 Static deformation density map in the purine plane of adenosine, based, as far as we are aware, on the fastest data set ever acquired for electron-density work (exposure time $<$ 1 h) [15] Copyright 2005 and Reproduction with Permission from American Chemistry Society.

The conclusion reached from all the experimental conditions discussed above is that if beam conditions are stable synchrotron radiation is the first choice for high-resolution electron-density data collection. This is especially true if smaller crystals are to be used; further advantages are reduced absorption and extinction. Cooling to low temperatures of approximately 10–20 K is superior to the 100 K cooling normally used, because many significant high-order reflections can be measured. If, however, crystals are large enough and diffract properly, MoKα radiation can be sufficient, as was demonstrated by the data collection of vitamin B_{12} at 100 K.

The 12-h sequence of diffraction experiments described above suggests the possibility of establishing high-throughput techniques in electron density research, thereby making electronic information of entire classes of chemically or biologically related compounds available at an increased pace. Together with the invariom data base approach that will be detailed in the next sections, high-speed evaluation of electron densities could become a routine task to be conducted in a time comparable with those currently needed for conventional X-ray analyses.

12.3
Studying Transferability with QTAIM – Atomic and Bond Topological Properties of Amino Acids and Oligopeptides

A key feature of Bader's theory of atoms in molecules is the partitioning of a molecular structure into sub-molecular regions, functional groups, or single atoms. The partitioning procedure makes use of the zero-flux surface (ZFS) in the electron-density gradient vector field $\nabla\rho(\mathbf{r})$. Surfaces of this type establish atomic basins around nuclear attractors of the corresponding trajectories of $\nabla\rho(\mathbf{r})$ and

Table 12.1 Summary of available experimental charge-density studies on amino acids (from [19]).

Compound	Radiation (Å)	Res. ($Å^{-1}$)	Temp. (K)	Refs
α-Gly	(X–N)	1.2	120	20
α-Gly	Mo, 0.71	1.15	23	21
L-Ala	Mo, 0.71	1.08	23	22, 23
DL-Val	Sy, 0.45	1.54	100	24, 25
DL-Ser	Mo, 0.71	1.22	123	26
DL-Ser	Sy, 0.45	1.54	100	26, 24, 25
L-Thr	Ag, 0.56	1.34	19	26
L-Cystine	Mo, 0.71	1.123	110	8
L-Asn.H_2O	Sy, 0.64	1.073	20	27
L-Asn	Sy, 0.53	1.46	100	26, 24, 25
L-Gln	Mo, 0.71	1.08	130	28
DL-Asp	Ag, 0.56	1.368	20	29
DL-Glu	Sy, 0.53	1.3	100	26, 24, 25
DL-Lys.HCl	Sy, 0.5	1.38	100	24, 25
DL-Arg.H_2O	Sy, 0.49	1.4	100	30
DL-His	Mo, 0.71	1.23	110	31
L-Trp.HCO_2H	Sy, 0.54	1.38	100	32
DL-Pro.H_2O	Sy, 0.50	1.12	100	6, 24, 25
L-Phe.HCO_2H	Mo, 0.71	1.18	25	19

uniquely define atomic volumes. Together with identification of critical points on bond paths, rings, and cages, tools are at hand for quantitative evaluation of bonding, atomic, or functional group properties. Transferability of the above-mentioned quantitative data is essential for application of database approaches to modeling of the electron density of larger systems. On the electronic level, it can be expected that density and derived properties of a functional group composed of atomic fragments should have high transferability when compared for different but chemically related molecules.

The biologically important class of the twenty genetically encoded amino acids was one of the first in which this transferability was systematically examined experimentally and theoretically. Bader and Matta have published complete topological data on all twenty amino acids based on theoretical calculations [16–18] and experimental studies on sixteen of the twenty amino acids have been performed by different groups, as detailed in Table 12.1. This class of compounds is thus the first for which a complete set of theoretical electron-density data is available and for which the corresponding experimental studies are approaching completeness.

Quantitative results for bond topological properties are summarized in Fig. 12.6 and Table 12.2 [19]. Figure 12.6a shows $\rho(\mathbf{r}_{bcp})$ values and Laplacians for the five main chain bonds common to all amino acids (CO, CO(H), CN, C_α–C′ (= CCO) and C_α–C_β (= CCR)) from the sixteen experimental studies. With the

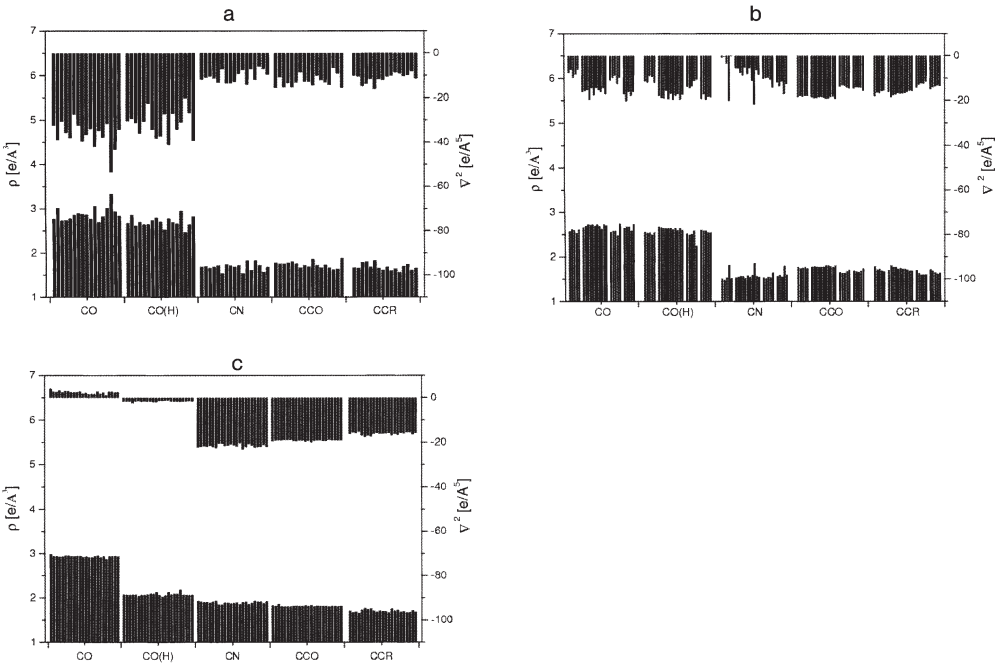


Fig. 12.6 Representation of the $\rho(\mathbf{r}bcp)$s and Laplacians of the five main chain bonds in amino acids [19] Copyright 2006 and Reproduction with Permission from Oldenbourg Publisher, Germany. (a) Experimental results. (b) Data from theoretical calculations at experimental geometry. Block separation according to different basis sets was used: HF/6-311++G**, HF/6-311++G(3df,3pd), B3LYP/6-311++G** and B3LYP/6-311++G(3df,3pd), respectively. (c) Data from ab initio calculations for neutral amino acids (Matta and Bader, RHF/6-311++G** [18]).

exception of L-phenylalanine, which contains a neutral carboxyl group, all experimental results are for negatively charged carboxylate groups. In all cases both C′–O bond lengths differ mainly because of weak intermolecular interactions, so the $\rho(\mathbf{r}_{\mathrm{bcp}})$ values for these bonds are not merged. CO(H) stands for the longer C′–O bond, which is commonly related to a lower $\rho(\mathbf{r}_{\mathrm{bcp}})$ value.

The corresponding results from calculation of experimental geometries are listed in Fig. 12.6b. Different authors have used different basis sets. Because, as discussed elsewhere [24], the effect of the wavefunctions used is not negligible, the entries are separated into four blocks according to the basis sets used.

Matta's and Bader's theoretical results for the 20 amino acids are listed in Fig. 12.6c. For their calculations they separated the amino acid molecules into the main-chain part, which is set to neutral, and the side-chain part, which is supposed to be the dominant state at pH 7; this implies that Lys, Arg, and His are protonated and the alkyl side-chains of Asp and Glu are deprotonated. The diagram illustrates the different topological properties of the two different CO bonds in neutral amino acids compared with the zwitterionic forms in Figs 12.6a and

Table 12.2 Mean values and absolute and relative standard deviations (σ and σ/ρ) for the five main chain bonds. First line – experiment (multipole); second–fifth line – calculations at experimental geometry (HF/6-311++G**; HF/6-311++G(3df,3pd); B3LYP/6-311++G**; B3LYP/6-311++G(3df,3pd)); sixth line – theoretical values (Matta and Bader, RHF/6-311++G**) [19].

Bond	ρ(r) (e Å^{-3})	σ (e Å^{-3})	σ/ρ(r) (%)	n[a]	$\nabla^2\rho$(r) (e Å^{-5})	σ (e Å^{-5})	$\sigma/\nabla^2\rho$(r) (%)	n[a]
CO	2.857	0.110	3.8	16	−35.631	4.263	12.0	16
	2.590	0.037	1.4	5	−7.540	1.557	20.6	5
	2.702	0.033	1.2	11	−16.045	1.625	10.1	11
	2.594	0.096	3.7	5	−10.326	1.430	13.9	5
	2.668	0.055	2.0	5	−17.470	1.728	9.9	5
	2.938	0.021	0.7	24	−2.480	0.618	24.9	24
CO(H)	2.717	0.105	3.9	16	32.356	5.170	16.0	16
	2.540	0.023	0.9	5	10.680	1.474	13.8	5
	2.641	0.023	0.9	11	18.138	1.331	7.3	11
	2.533	0.041	1.6	4	13.248	1.431	10.7	4
	2.578	0.027	1.0	5	18.590	0.988	5.3	5
	2.081	0.033	1.6	24	1.680	0.226	13.5	24
CN	1.685	0.083	4.9	17	10.506	2.652	25.2	17
	1.547	0.139	8.8	5	5.020	8.579	170.9	5
	1.586	0.093	5.9	11	8.540	4.754	55.7	11
	1.538	0.013	0.8	4	10.120	0.426	4.2	4
	1.622	0.101	6.2	5	13.320	2.326	17.5	5
	1.896	0.027	1.4	24	21.739	0.711	3.3	24
CCO	1.735	0.072	4.2	17	12.594	2.657	21.1	17
	1.758	0.013	0.7	5	18.120	0.268	1.5	5
	1.787	0.016	0.9	11	18.830	0.398	2.1	11
	1.665	0.028	1.7	5	13.850	0.588	4.2	5
	1.688	0.031	1.8	5	14.528	0.602	4.1	5
	1.817	0.013	0.7	24	19.247	0.273	1.4	24
CCR	1.681	0.076	4.5	16	11.231	2.136	19.0	16
	1.718	0.046	2.7	5	16.520	0.988	6.0	5
	1.735	0.038	2.2	11	16.822	0.842	5.0	11
	1.624	0.045	2.8	5	12.744	1.023	8.0	5
	1.664	0.040	2.4	5	13.942	0.713	5.1	5
	1.705	0.031	1.8	24	16.260	0.606	3.7	24

[a] n is the number of contributing entries.

12.6b. The positive Laplacians of the CO bonds describe partial ionic character. As a result of this the $\rho(\mathbf{r}_{bcp})$ values of the CN bond are higher than for the zwitterionic molecules, in accordance with chemical understanding. In the experimental data there is a small difference between the C′–O bonds which originates in the weak interactions described.

Averages and statistical standard deviations (σ values) of the quantities illustrated in Fig. 12.6 are listed in Table 12.2. For the experimental averages of the five bonds considered (first lines in Table 12.2) the σ values are, approximately, within 0.07–0.11 e Å^{-3} and 2–5 e Å^{-5} for $\rho(\mathbf{r}_{bcp})$ and $\nabla^2\rho(\mathbf{r}_{bcp})$, respectively. For the 24 entries of Matta and Bader's theoretical calculations (all obtained with the same basis set RHF/6-311++G**, last lines for each bond in Table 12.2) the σ values are much smaller, and here intermolecular interactions are not considered. It was, nevertheless, shown that results of *ab initio* calculations with different basis sets can vary in the same range as experimental data.

It can be concluded that despite a variety of experimental and refinement conditions and different β-substituents and crystal environments, the overall results for experiment and theory are very consistent in their respective ranges. With regard to Bader's concept of transferability of submolecular properties these data are useful for predicting the bond properties of larger systems of biological interest, which often cannot be determined experimentally.

Between single amino acids and macromolecules, for example proteins, are the oligopeptides, which contain the building blocks of proteins. We would like to compare the transferability of atomic volumes and charges in the peptide bond region for several oligopeptides. For this purpose we make use of experimental data for five dipeptides, one hexapeptide, and four tripeptides. The latter are part of a systematic study of tripeptides of the type L-ala–XXX–L-ala, were the central amino acid residue XXX was varied to examine whether this central residue affects transferability in the peptide bond region.

The results obtained so far are summarized in Fig. 12.7. The averages of comparable quantities show that the internal consistency for volumes is <1 Å^3. The atomic volumes at the other C_α and C' atoms are equal within twice the statistical error, except if C_α belongs to a Gly residue. The average volumes of the (non-Gly) C_α atoms are more than 1 Å^3 smaller than those of the Gly C_α atoms, in which the second hydrogen atom enables the carbon to expand. A nearest-neighbor effect is also seen for the N atoms. In proline the nitrogen atom is part of the five membered ring and bonded to a third carbon atom instead of hydrogen, which reduces the volume by more than 1 Å^3, by analogy with the above quoted volume expansion for the glycine C_α.

The QTAIM charges (for averages see also Fig. 12.7) agree within the given atom types by 0.07–0.16 e, which is a surprisingly small spread. The C_α atoms carry a small positive charge, the hydrogen atoms of the peptide N–H are moderately positively charged and the C' atoms carry a high positive charge, whereas strong negative charges close to 1 e are observed on the N and O atoms. These experimental results indicate that polarization of Bader atoms is much higher than obtained, for example, from theoretical orbital models (NBO or Mulliken charges) [35] or than used in force field parameterization. For example, the AMBER [36] force field uses charges of -0.5 for oxygen, $+0.5$ for carbon, -0.57 for nitrogen, and $+0.37$ e for hydrogen atoms in the peptide bond. The definition and determination of atomic charges have been subjects of controversial discussions in recent years [35] and charges derived from different methods may differ significantly.

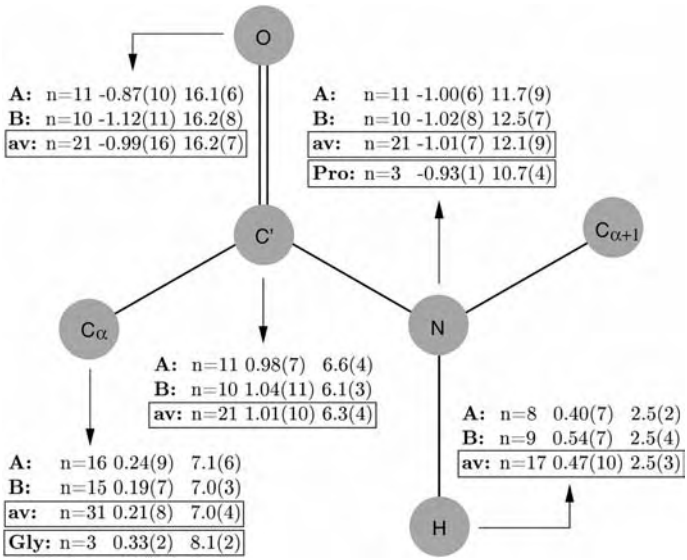


Fig. 12.7 Average atomic charges (e) and volumes (Å^3) of the atoms in the peptide bond region. A refers to five dipeptides and one hexapeptide [33]; B refers to the four tripeptides AAA.$\frac{1}{2}H_2O$ [34], APA.H_2O, AYA.H_2O, and AYA.C_2H_5OH (to be published); av is the average over all entries. n is the number of contributing entries.

QTAIM charges are based on well-defined atomic segments of the electronic charge density and can be derived from an experiment [37, 38].

The positive charges on the C_α, C′, and H atoms total approximately $+1.7\ e$ and the negative charges on N and O amount to approximately $-2e$, so for each peptide bond region an excess of $-0.3\ e$ must be compensated by the side-chains or, for glycine, another hydrogen atom.

A preliminary conclusion of this study of the peptide bond is that very reproducible atomic properties for the contributing atoms can be derived if the chemical environment is comparable. A significant experimentally detectable effect of next-nearest neighbors on the electron density of C_α-type atoms was not observed. The results therefore show the validity of the nearest/next-nearest neighbor approximation and encourage the use of database approaches for electron-density modeling of macromolecules.

12.4 Invariom Modeling

Conventional interpretation of atomic-resolution X-ray data is based on the promolecule model (superimposition of spherically symmetric, isolated atomic den-

sities), also termed the independent atom model (IAM). At this level the exactness of the reconstruction of the diffraction image is limited to the main characteristic of the topology of the crystalline electron-density distribution, that is, the occurrence of maxima at the nuclear positions.

To replace the IAM with an aspherical scattering model we recently introduced the concept of invarioms (pseudoatoms that are invariant on transfer from one molecule to another) [39]. A pseudoatom is a density fragment of the multipole model (rigid pseudoatom formalism) that was introduced in 1978 [4] on the basis of earlier work [40], and has been described in full detail elsewhere [41, 42]. ''Individual'' invariom aspherical scattering factors take into account the chemical environment of a bonded atom. A finite set of invarioms, identified by connectivity and bonding, can be assigned to each chemical element. The density of an invariom is extracted from a quantum chemical geometry optimization of a small model compound containing the invariom and its defining nearest/next-nearest neighbor atoms terminated by hydrogen atoms or a suitable fragment in the second coordination sphere, as will be detailed later. For the purpose of building a database of such invarioms, theoretical structure factors [43] were calculated from geometry-optimized model compounds. These ''simulated data'' are transformed by least-squares refinement into multipole electron-density data for the invariom. The total density of a molecule is obtained by superimposition of its constituting invarioms. When used to interpret conventional X-ray data of moderate resolution, complex invariom scattering factors improve structure refinement almost to the extent of a free multipole refinement with high-resolution data, and the geometry obtained by invariom structure refinement of low-resolution data is almost identical as when high-resolution data are used. Figure 12.8 illustrates selection of invarioms and their associated model compounds for the structure of DL-serine. The scheme also shows the invariom nomenclature that will be explained later.

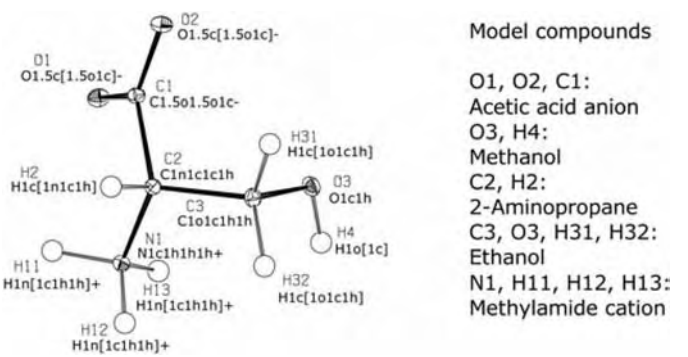


Fig. 12.8 ORTEP plot [44] of the molecular structure of DL-serine at 20 K with atomic numbering scheme, invarioms assigned using notation explained in Section 12.4.1, and the model compounds used to obtain the invariom electron density.

The method described here differs in important aspects from the experimental approach suggested by Pichon-Pesme et al. [45–47] for ultra-high resolution protein crystallography. Because our approach does not rely on experimental data, measurement errors and bias arising as a result of inadequate modeling of the thermally smeared density are excluded. We also provide a consistent definition for transferable sites in terms of invarioms that can be generated, in principle, for all chemically relevant bonding situations at any level of theory.

Some relevant details of a similar database approach by Volkov et al. [48] must be mentioned. In the work of Volkov et al. populations are averaged over model compounds that were chosen from a large possible range, so that deviation from electrical neutrality is different from results using our invariom database, in which only one defined model compound is used for each invariom. Rules for setting up a model compound will be given below. Inclusion of a notation scheme in the invariom concept also enables automation of least-squares refinement of experimental X-ray data.

12.4.1 Invariom Notation, Choice of Model Compounds, and Practical Considerations

In invariom notation the element symbol of the atom of interest, in capitals, begins the name, then formal bond order and nearest neighbors follow in lower case ordered by their position in the periodic table (heavier atoms first). The order of the ligands in the invariom name is determined by their decreasing bond order. For mesomerism or delocalized systems a bond order of 1.5 is specified. For these, the next-nearest neighbors must be taken into account in the name. Next-nearest neighbors are also used for invariom names of hydrogen and hypervalent atoms. This is achieved by writing the next-nearest neighbors according to the previous rules in brackets after the nearest neighbors (examples are given in Fig. 12.8). For chiral invarioms, *R* or *S* is used as a prefix (CIP rules), separated by a dash. Because, usually, next-nearest neighbors only are considered, chiral invarioms occur a lot less frequently than chiral atoms. Finally a + or a − indicates a charge when the model compound used is an ion, for example the acetic acid anion. The sign is best placed at the very end of the name.

The invariom approach differentiates between chemical environments by distinguishing between single and mesomeric bonds, which enables rules to be set up for generation of model compounds. For single-bond systems the model compounds simply include the nearest neighbors saturated by hydrogen atoms. This is also valid for double/triple bonded systems. Charged groups, for example $R–NH_3^+$, require a charged model compound. For hydrogen atoms, next-nearest neighbors must be included to achieve proper electroneutrality during summation over invariom fragments at the application stage. Next-nearest neighbors are also included for delocalized systems and hypervalent atoms. When extended mesomeric or delocalized systems occur in a structure, the whole mesomeric fraction of the molecule must be included in the model compound in a suitable

manner. A good example is the model compound of a carbon invariom assigned to any atom of a benzene ring. The obvious advantage of calculating the complete ring is that use of a radical – here the propenyl-radical taking into account nearest neighbors only – can be avoided.

For practical application of invariom modeling analysis of the IAM geometry is the starting point. For this purpose the empirical relationship below [49] is used to identify covalent bonds, where d is the distance between the atoms, in Ångstroms, r_c is the covalent radius, and $|\Delta(EN)|$ the difference between the Allred–Rochow electronegativities [50] of two atoms in a structure.

$$d \cdot 0.85 \leq r_c(atom1) + r_c(atom2) - 0.08|\Delta(EN)| \tag{1}$$

When $d \cdot 0.85$ is smaller than the sum of the covalent radii of the two atoms multiplied by an electronegativity difference a bond is found. To differentiate between single/mesomeric/double, and triple bonds the same empirical relationship, Eq. (1), is used to assign a bond distinguishing term χ (Eq. 2):

$$\chi = [r_c(atom1) + r_c(atom2) - 0.08|\Delta(EN)|] - d \tag{2}$$

The distinction between a singe/mesomeric/double bond is arbitrary and not very distinctive for some delocalized systems, but experience with a large number of structures and optimized model compounds led to the establishment of reliable values for sensible distinction where to include next-nearest neighbors and where it is unnecessary.

Software [51] has been developed that enables the automation of the modeling process. It automatically analyses molecular geometry, assigns invariom names to each atom in a structure, uses invariom notation to find and transfer database density values and writes input files for the respective aspherical-atom refinement program XD [52].

12.4.2 Support for Pseudoatom Fragments from QTAIM

We were interested whether or not it is possible to support the invariom approach by means of QTAIM. Integration over the atomic basin as defined by the zero flux surface of $\rho(\mathbf{r})$ yields well-defined atomic charges, and we will focus on theoretically calculated AIM charges of several chemically related example molecules to investigate the limits of transferability.

The level of transferability needed for the X-ray scattering model is limited to the possible accuracy and resolution of the experimental data. Although perfect transferability is not reached in theory, current experimental limitations enable approximations to be made. One factor reducing transferability is hydrogen bonding; different conformations [16] can also lead to slightly different charges.

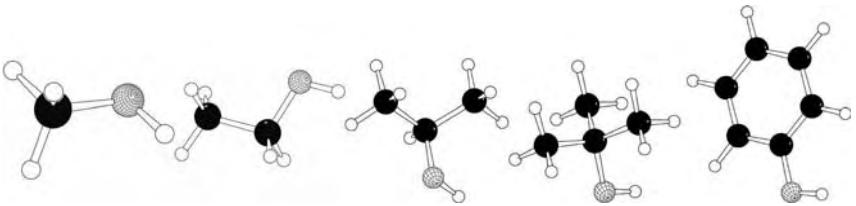


Fig. 12.9 Series of alcohols studied with high OH fragment transferability – from left to right methanol, ethanol, isopropanol, 2-butanol, and phenol. On the other hand, the charge on the carbon atom to which the oxygen is attached differs by a large amount. This is readily explained by the different chemical environments of the carbon atoms. Whereas in methanol only three hydrogen atoms are attached to it, in ethanol one, in isopropanol two, and in isobutanol three methyl groups are also found with the corresponding hydrogen atoms.

Bader et al. have shown for a series of hydrocarbons and silanes that a CH_2 group contributes an identical amount to the energy [53]. For other, more complicated fragments, assuming QTAIM transferability is less valid and the approach to subdivide a molecule into QTAIM atoms and functional groups less successful. While application is justified for classes of molecules with identical fragments, this approach is not suited to a generalized scattering model. The choice of an atom-based fragment is, furthermore, left to the researcher.

To illustrate these points several compounds were geometry-optimized using the basis set D95++(3df,3pd) as available in the software GAUSSIAN [54]. Their electron densities were then analyzed with Bader's QTAIM (MORPHY software [55]).

In the first series of alcohols shown in Fig. 12.9, the oxygen atom is always bonded to a carbon atom and a hydrogen atom. The atomic charge of the oxygen atom is almost identical in these example molecules, and for the hydrogen atom attached to it (Table 12.3). Small differences are because of intermolecular hydrogen bonding of the oxygen atom and due to different conformations. On can conclude the OH group has very high transferability.

Table 12.3 Atomic charges for a series of alcohols.

Molecule	Atom	Charge (*e*)	Atom	Charge (*e*)	Atom	Charge (*e*)
Methanol	O	−1.10	H	0.56	C	0.61
Ethanol	O	−1.10	H	0.56	C	0.59
Isopropanol	O	−1.10	H	0.55	C	0.56
2-Butanol	O	−1.10	H	0.55	C	0.53
Phenol	O	−1.12	H	0.58	C	0.48

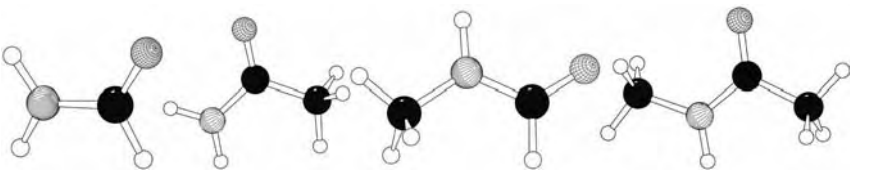


Fig. 12.10 Series of amides with low fragment transferability. From left to right: formamide, acetamide, *N*-methylformamide, and *N*-methylacetamide.

If one wishes to reconstruct the electron density based on the sum of atomic QTAIM fragments as successfully attempted earlier [56], the carbon atom attached to the OH group cannot be regarded as transferable and the assumption of carbon transferability would introduce a systematic error in a scattering model, whereas the oxygen and hydrogen atoms have very similar or identical charges.

We have also calculated atomic QTAIM charges of a number of amides (Fig. 12.10 and Table 12.4). For these molecules transferability is only found for the oxygen atoms that are partly double-bonded to the carbon atom of the peptide bond. The nitrogen atoms of the amino group are similar in formamide and acetamide, and in *N*-methylformamide and *N*-methylacetamide, but differ in the two groups. The situation is analogous with the carbon atom – charges differ between substituted and nonsubstituted atoms, but are in agreement when the neighbors are similar. We conclude that, to a good approximation, atoms have similar QTAIM charges when their nearest neighbored atoms are identical [57]; this approach is termed the nearest-neighbor approximation (NNA).

We can also observe "compensatory transferability" in the charges, to use the term introduced by Bader [53]. When two possible mesomeric Lewis formulae can be written for a molecule, charge compensation can be assumed to be more pronounced than for a molecule consisting of single-bonded atoms only.

These observations illustrate the reasons for the choice of model compounds used to predict the electron density for invariom modeling as mentioned above.

Table 12.4 Atomic charges for a series of amides.

Molecule	Atom	Charge (*e*)	Atom	Charge (*e*)	Atom	Charge (*e*)
Formamide	O	−1.16	N	−1.16	C	1.50
Acetamide	O	−1.18	N	−1.16	C	1.44
N-methylformamide	O	−1.16	N	−1.15	C	1.49
N-methylacetamide	O	−1.17	N	−1.15	C	1.43

12.5
Applications of Aspherical Invariom Scattering Factors

12.5.1
Molecular Geometry and Anisotropic Displacement Properties

Invariom modeling improves the accuracy of molecular geometry from X-ray single crystal diffraction, especially for hydrogen atoms, for which bond distances become comparable with results from neutron diffraction or theoretical calculations. Modeling with aspherical scattering factors has been shown to be more appropriate than application of the IAM [58]. Asphericity shifts [59] are bond-length aberrations because of the use of the IAM. They occur, for example, in longer C–O bonds, because of spherical averaging over the oxygen valence density in the IAM at low resolution; C=C bonds can also be affected. Asphericity shifts disappear in invariom model refinements, although such differences are often within the error range of the least squares.

Improvement of molecular geometry for DL-serine by using invarioms is shown in Table 12.5 by comparison of bond distances of the 298 K X-ray invariom data with room-temperature neutron study results for different resolution [60]. In essence, invariom modeling provides an improved scattering factor model and, as a result, the least-squares fit of aspherical scatterers to experimental structure factors usually leads to substantial improvement of the crystallographic *R*-factors and the goodness of fit, and to a reduction of the remaining residual electron density.

Table 12.5 Neutron (N) bond lengths of DL-serine compared with X-ray (X) invariom and the IAM model at room temperature.

Bond	Distance (N)	Distance (X, invarioms)		Distance (X, IAM)	
		$\sin\theta/\lambda = 0.98$	$\sin\theta/\lambda = 0.6$	$\sin\theta/\lambda = 0.98$	$\sin\theta/\lambda = 0.6$
O(1)–C(1)	1.248(1)	1.2487(5)	1.248(2)	1.2479(7)	1.246(2)
O(2)–C(1)	1.257(1)	1.2585(5)	1.257(2)	1.2586(8)	1.256(2)
O(3)–C(3)	1.414(1)	1.4150(5)	1.414(2)	1.4172(8)	1.417(2)
O(3)–H(4)	0.981(1)	0.95(1)	0.97(2)	0.92(2)	0.92(3)
N(1)–C(2)	1.487(1)	1.4884(5)	1.488(2)	1.4883(7)	1.490(2)
N(1)–H(11)	1.037(1)	1.048(9)	1.05(2)	0.96(2)	0.94(2)
N(1)–H(12)	1.045(1)	1.03(1)	1.03(2)	0.95(2)	0.96(2)
N(1)–H(13)	1.041(1)	1.03(1)	1.02(2)	0.94(2)	0.94(2)
C(1)–C(2)	1.531(1)	1.5316(5)	1.532(2)	1.5291(7)	1.528(2)
C(2)–C(3)	1.518(1)	1.5201(5)	1.519(2)	1.5165(8)	1.514(2)
C(2)–H(2)	1.101(1)	1.080(7)	1.07(2)	0.956(9)	0.93(2)
C(3)–H(31)	1.095(1)	1.096(9)	1.10(2)	0.97(2)	0.97(2)
C(3)–H(32)	1.095(1)	1.121(9)	1.12(2)	1.00(2)	0.98(2)

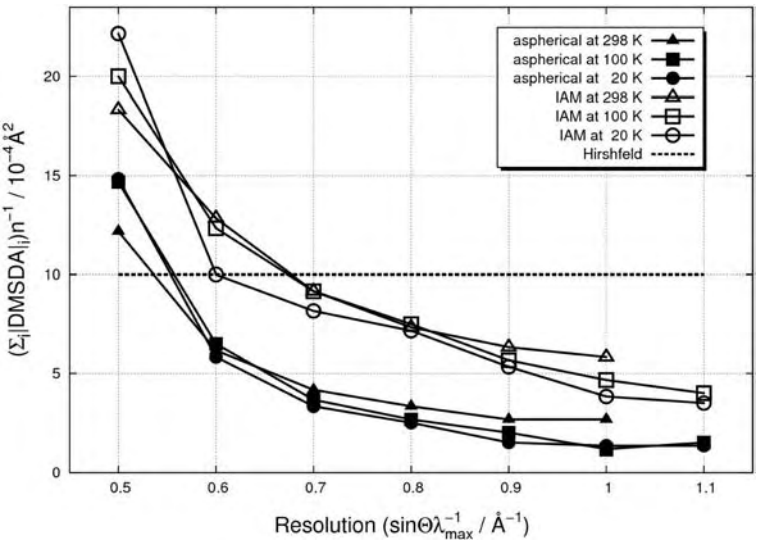


Fig. 12.11 Mean value for bonds between C, N, and O for the Hirshfeld test plotted against resolution. Above 10×10^{-4} Å^2 (dotted line) the test is not regarded as fulfilled, as suggested by Hirshfeld [58] Copyright 2005 and Reproduction with Permission from IUCr.

Another result of the modeling process is the increased physical significance of the anisotropic displacement parameters that describe thermal motion (and disorder) in a structure, as can be proven by the results of the Hirshfeld test [5]. Figure 12.11 shows the effect of the inclusion of the aspherical density in the anisotropic temperature parameters with regard to data resolution for DL-serine. Temperature data for carbon, nitrogen and oxygen are not regarded as including bonding effects if the difference of the mean-square displacement amplitudes (DMSDA) is smaller than 0.001 Å^2. A mean DMSDA value for the six nonhydrogen bonds was used in Fig. 12.11 in the study of DL-serine [58] and this value was plotted against resolution for invariom scattering factors obtained with the basis set B3LYP/6-311++G(3df,3pd). Whereas for the IAM at 100 K deconvolution of electron density and thermal effects was not achieved and the test was not fulfilled at a resolution of 0.55 Å^{-1} in $\sin\theta/\lambda$ (or $d = 0.9$ Å), by using the invariom model the Hirshfeld test was fulfilled for all three temperatures (20 K, 100 K, 298 K) investigated. It was concluded that this, or higher, resolution is recommended for invariom modeling.

12.5.2
Using the Enhanced Multipole Model Anomalous Dispersion Signal

Absolute configuration of light atom structures is of crucial importance in the pharmaceutical industry. Although the phenomenon of chirality was discovered many years ago, the relevance of the absolute structure of a drug applied to the

human organism was not recognized until the Contergan/Thalidomide scandal in the early sixties. These events made clear that absolute structure can be of outmost importance and that it should be carefully analyzed for every drug candidate before it is registered as a drug, even though it was later found that Thalidomide racemizes *in vivo*. The proportion of single-enantiomer drugs among new drugs introduced into the market is rapidly growing, and reached ~70% in 2002 [61]. There are stringent legal requirements to determine the absolute configuration of drug molecules. A related aspect is the possibility of extending or sidestepping patents by invoking chirality. Single-crystal X-ray structure analysis enables determination of the absolute structure from the intensity difference between Friedel pairs, because of anomalous dispersion. The objective of such an analysis is, usually, to assign absolute configuration to chiral molecules in a crystal structure and considerable effort has been devoted to this research topic. Introduction of the x parameter by Flack [62] enabled unambiguous assignment of enantiomorph polarity and the x parameter is widely used and implemented in many least-squares refinement software, for example CRYSTALS [63], SHELXL [64] or GFMLX [65]. Absolute determination of structure requires that the anomalous dispersion signal is sufficiently pronounced. We have recently shown that the multipole model provides additional inversion distinguishing power [66]. Invariom modeling with fixed density terms provides a means of exploiting this to improve absolute structure determination with CuKα, and it has been found that high-resolution MoKα datasets can provide almost similar information when evaluated with the invariom approach.

12.5.3
Modeling the Electron Density of Oligopeptide and Protein Molecules

For the purpose of structural refinement of amino acid, oligopeptide and protein molecules the naturally occurring amino acids were analyzed in terms of their invariom fragments, and a database with 73 entries that covers this class of compound was generated from 37 model compounds. This invariom database has been validated on 42 experimental small-molecule crystal structures (from IUCr journals), of different quality and resolution, covering not only the naturally occurring amino acids, but some of their derivatives, their protonated/unprotonated states, and most common solvents (details of the structures studied have been published elsewhere [67]).

Figure 12.12a shows the crystallographic R-factor for IAM and invariom models for these trial structures. $R(F)$ is always equal to or better for the invariom model, when compared with the IAM. The difference $R(F_{IAM}) - R(F_{inv})$ depends on the resolution and temperature of an experiment. Improvements are more apparent at lower temperatures, because of better deconvolution of thermal and electronic effects. Because low-resolution datasets do not contain as much information on the aspherical part of $\rho(\mathbf{r})$, improvement of the R-factor is smaller than when more extended datasets are used. Figure 12.12b compares the positive residual electron density $\Delta\rho(\mathbf{r})$ for the IAM and invariom models. Here, the most interesting feature can be observed for a high resolution structure containing sulfur,

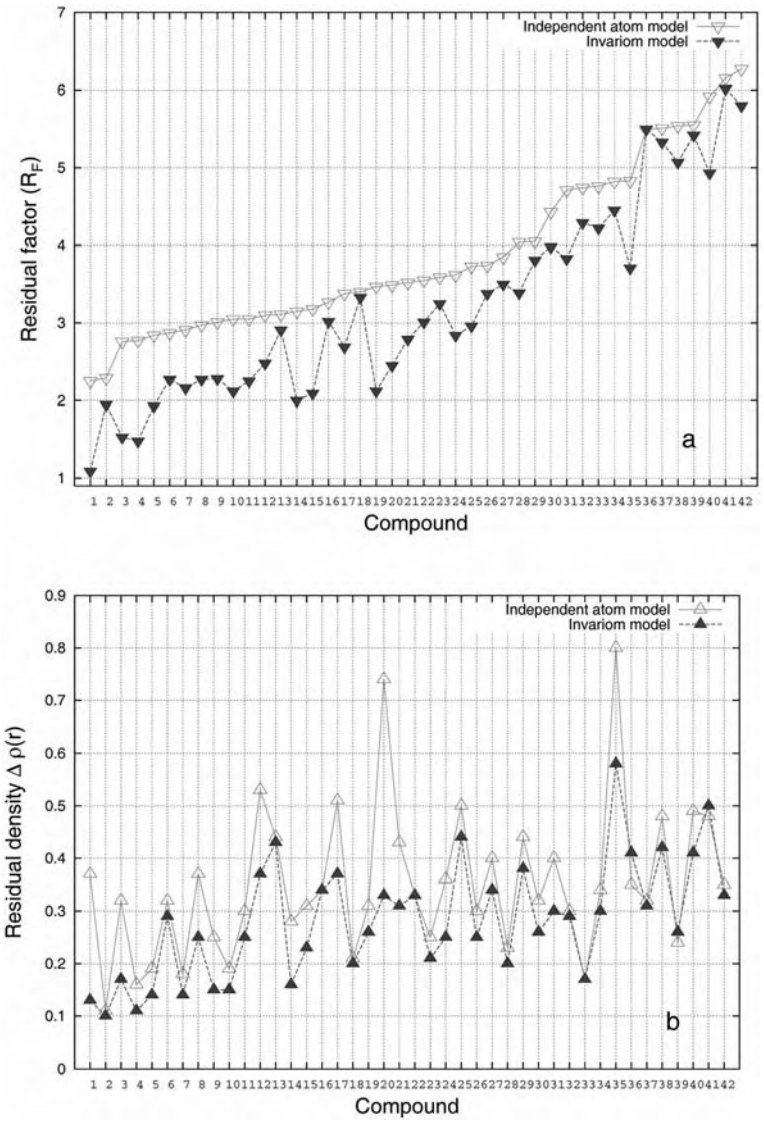


Fig. 12.12 (a) Comparison of the crystallographic R(F) factor for the IAM and invariom models. (b) Comparison of the residual density $\Delta\rho(\mathbf{r})$ for the IAM and invariom models [67] Copyright 2006 and Reproduction with Permission from IUCr.

where the residual density for the invariom model is substantially reduced and becomes comparable with that of other structures that do not contain sulfur.

In disordered structures analogous behavior is seen for $\Delta\rho(\mathbf{r})$ and reduction of the R-factor. For such structures the residual density is similar or even increased

by use of the invariom model. One can conclude that the remaining, unmodeled, density is the reason for only small improvements of the *R*-factor for disordered structures.

An important conclusion for protein refinement can therefore be drawn from disordered structures – it is the modeling of disorder and the completeness of the structural model, rather than the aspherical electron density contribution, that limit the fit of calculated and experimental structure factors and, therefore, the quality of the results. As already remarked, invariom modeling requires data resolution of $d \leq 0.9\,\text{Å}$ or $\sin\theta/\lambda \geq 0.55\,\text{Å}^{-1}$. This resolution is also recommended for protein data.

12.6 Conclusion

From systematic study of amino acids and oligopeptides by experimental electron density analysis using Bader's QTAIM an understanding of fragment transferability emerged. Use of appropriate model compounds now enables reproduction of molecular electron density $\rho(\mathbf{r})$ from fragments; the result is a good approximation to the total density when the invariom pseudoatoms density description is employed. A pseudoatomic fragmentation of a molecular electron density retaining the local chemical environment of an atom enables convenient extension of the IAM scattering model. Such invariom modeling improves the inversion distinguishing power in absolute structure determination and yields "charge-density quality" geometries from low-resolution standard X-ray datasets. Anisotropic displacement parameters become physically meaningful and all properties that can be derived from a model electron density can be rapidly calculated for larger molecules so that a convenient approach to electron densities of macromolecules (proteins, polynucleotides) is feasible.

To apply invariom modeling for standard small-molecule structures no further calculations nor extra experimental procedures are necessary, making it a rapid, readily accessible, and useful tool for standard crystallographic work. In this way high-throughput techniques, which have been established in a variety of fields of X-ray diffraction-based structure research, can also be applied in experimental electron-density work. This would be of special importance in the biological sciences. In medical chemistry, where it is a fundamental challenge to understand drug–target recognition processes, knowledge of electron-density distribution is a valuable completion of structure information and serves as a basis for a better understanding of such interactions than consideration of steric properties only. Since rapid screening of a large number of chemical compounds is indispensable in structure-guided drug discovery, the generation of electronic information of entire classes of chemically or pharmacologically related compounds would be highly desirable. This can become a routine task to be performed in time periods comparable to those currently needed for conventional X-ray analyses.

References

1 Debye, P. *Annalen der Physik* **46** (1915), 809.
2 Coppens, P. *Angew. Chem. Int. Ed.* **44** (2005), 6810.
3 Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*, Clarendon Press, Oxford, 1990.
4 Hansen, N. K., and Coppens, P. *Acta Cryst. A* **34** (1978), 909.
5 Hirshfeld, F. L. *Acta Cryst. A* **32** (1976), 239.
6 Koritsánszky, T., Flaig, R., Zobel, D., Krane, H.-G., Morgenroth, W., and Luger, P. *Science* **279** (1998), 356.
7 Iversen, B. B., Larsen, F. K., Pinkerton, A. A., Martin, A., Darovsky, A., and Reynolds, P. A. *Acta Cryst. B* **55** (1999), 363.
8 Dahaoui, S., Jelsch, C., Howard, J. A. K., and Lecomte, C. *Acta Cryst. B* **55** (1999), 226.
9 Wagner, A., Flaig, R., Zobel, D., Dittrich, B., Bombicz, P., Strümpel, M., Luger, P., Koritsánszky, T., and Krane, H.-G. *J. Phys. Chem. A* **106** (2002), 6581.
10 Lamparth, I., Maichle-Mössmer, C., and Hirsch, A. *Angew. Chem. Int. Ed.* **34** (1995), 1607.
11 Luger, P., Messerschmidt, M., Scheins, S., and Wagner, A. *Acta Cryst. A* **60** (2004), 390.
12 Dittrich, B., Koritsánszky, T., Volkov, A., Mebs, S., Luger, P., *Angew. Chem. Int. Ed.* **46** (2007), in press.
13 Zobel, D., Luger, P., Dreissig, W., and Koritsánszky, T. *Acta Cryst. B* **48** (1992), 837.
14 Messerschmidt, M., Meyer, M., and Luger, P. *J. Appl. Cryst.* **36** (2003), 1452.
15 Luger, P., Wagner, A., Hübschle, C. B., and Troyanov, S. I. *J. Phys. Chem. A* **109** (2005), 10177.
16 Matta, C. F., and Bader, R. F. W. *Proteins: Structure, Function and Genetics* **40** (2000), 310.
17 Matta, C. F., and Bader, R. F. W. *Proteins: Structure, Function and Genetics* **48** (2002), 519.
18 Matta, C. F., and Bader, R. F. W. *Proteins: Structure, Function and Genetics* **52** (2003), 360.
19 Mebs, S., Messerschmidt, M., Luger, P. *Z. Kristallogr.* **221** (2006), 656.
20 Legros, J.-P.; Kvick, A. *Acta Cryst. B* **36** (1980), 3052.
21 Destro, R., Roversi, P., Barzaghi, M., and Marsh, R. E. *J. Phys. Chem. A* **104** (2000), 1047.
22 Destro, R., Marsh, R. E., and Bianchi, R. *J. Phys. Chem.* **92**, 4 (1988), 966.
23 Gatti, C., Bianchi, R., Destro, R., and Merati, F. *J. Mol. Struct. (Theochem)* **255** (1992), 409.
24 Flaig, R., Koritsanszky, T., Dittrich, B., Wagner, A., and Luger, P. *J. Am. Chem. Soc.* **124** (2002), 3407.
25 Flaig, R. *Neue experimentelle Methoden der Ladungsdichtebestimmung*. PhD thesis, Freie Universität Berlin, 2000.
26 Flaig, R., Koritsánszky, T., Janczak, J., Krane, H.-G., Morgenroth, W., and Luger, P. *Angew. Chem. Int. Ed.* **38**, 10 (1999), 1397.
27 Arnold, W. D., Sanders, L. K., McMahon, M. T., Volkov, A. V., Wu, G., Coppens, P., Wilson, S. R., Godbout, N., and Oldfield, E. *J. Am. Chem. Soc.* **122** (2000), 4708.
28 Wagner, A., and Luger, P. *J. Mol. Struct.* **595** (2001), 39.
29 Flaig, R., Koritsánszky, T., Zobel, D., and Luger, P. *J. Am. Chem. Soc.* **120**, 10 (1998), 2227.
30 Kingsford-Adaboh, R., Dittrich, B., Wagner, A., Messerschmidt, M., Flaig, R., and Luger, P. *Z. Krist.* **217** (2002), 168.
31 Coppens, P., Abramov, Y., Carducci, M., Korjov, B., Novozhilova, I., Alhambra, C., and Pressprich, M. R. *J. Am. Chem. Soc.* **121**, 11 (1999), 2585.
32 Scheins, S., Dittrich, B., Messerschmidt, M., Paulmann, C., and Luger, P. *Acta Cryst. B* **60** (2004), 184.

33 Dittrich, B. *Herleitung atomarer Eigenschaften von Oligopeptiden aus ihren experimentellen Elektronendichten*. PhD thesis, Freie Universität Berlin, 2002.

34 Rödel, E., Messerschmidt, M., Dittrich, B., and Luger, P. *Org. Biomol. Chem.* **4** (2006), 475.

35 Wiberg, K. B., and Rablen, P. R. *J. Comp. Chem.* **14**, 12 (1993), 1504.

36 Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Jr., K. M. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. *J. Am. Chem. Soc.* **117** (1995), 5179.

37 Bader, R. F. W., and Zou, P. F. *Chem. Phys. Lett.* **191** (1992), 54.

38 Bader, R. F. W., and Matta, C. F. *The J. Phys. Chem. A* **108** (2004), 8385.

39 Dittrich, B., Koritsánszky, T., and Luger, P. *Angew. Chem. Int. Ed.* **43** (2004), 2718.

40 Stewart, R. F. *J. Chem. Phys.* **51**, 10 (1969), 4569.

41 Coppens, P. *Ann. Revi. Phys. Chem.* **43** (1992), 663.

42 Coppens, P. *X-Ray Charge Densities and Chemical Bonding*, Oxford University Press, Oxford, 1997.

43 Chandler, G. S., and Spackman, M. A. *Acta Cryst. A* **34**, 2 (1978), 341.

44 Burnett, M. N., and Johnson, C. K. ORTEP-III, Oak Ridge National Laboratory Report ORNL-6895, Oak Ridge, Tennessee, 1996.

45 Pichon-Pesme, V., Lecomte, C., and Lachekar, H. *J. Phys. Chem.* **99**, 16 (1995), 6242.

46 Jelsch, C., Pichon-Pesme, V., Lecomte, C., and Aubry, A. *Acta Cryst. D* **54** (1998), 1306.

47 Pichon-Pesme, V., Jelsch, C., Guillot, B., and Lecomte, C. *Acta Cryst. A* **60** (2004), 204.

48 Volkov, A., Li, X., Koritsánzky, T., and Coppens, P. *J. Phys. Chem. A* **108** (2004), 4283.

49 Schomaker, V., and Stevenson, D. P. *J. Am. Chem. Soc.* **63** (1941), 37.

50 Allred, A. L., and Rochow, E. G. *J. Inorg. Nucl. Chem.* **5** (1958), 264.

51 Hübschle, C. B., and Dittrich, B. INVARIOMTOOL, a Preprocessor Program for Aspherical Atom Modeling with XD Using Invarioms. Freie Universität Berlin, 2004.

52 Koritsánszky, T., Richter, T., Macchi, P., Volkov, A., Gatti, C., Howard, S., Mallinson, P. R., Farrugia, L., Su, Z. W., and Hansen, N. K. XD – a Computer Program Package for Multipole Refinement and Topological Analysis of Electron Densities from Diffraction Data. Freie Universität Berlin, 2003.

53 Bader, R. F. W., and Bayles, D. *J. Phys. Chem. A* **104** (2000), 5579.

54 Frisch, M. J., Trucks, G. W., Schlegel, H. B., G. E. Scuseria, M. A. R., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Jr., R. E. S., Burant, J. C., Dapprich, S., J. M. Millam, A. D. D., Kudin, K. N., Strain, M. C., Farkas, O., J. Tomasi, V. B., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., C. Adamo, S. C., Ochterski, J., Petersson, G. A., Ayala, P. Y., Q. Cui, K. M., Rega, N., Salvador, P., Dannenberg, J. J., D. K. Malick, A. D. R., Raghavachari, K., Foresman, J. B., J. Cioslowski, J. V. O., Baboul, A. G., Stefanov, B. B., Liu, G., A. Liashenko, P. P., Komaromi, I., Gomperts, R., Martin, R. L., D. J. Fox, T. K., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., M. Challacombe, P. M. W. G., Johnson, B., Chen, W., Wong, M. W., J. L. Andres, C. G., Head-Gordon, M., Replogle, E. S., and Pople, J. A. Gaussian 98, Revision A.11.3. Gaussian Inc., Pittsburgh PA, 2002.

55 Popelier, P. L. A., and Bone, R. G. A. MORPHY98, a Topological Analysis Program written by PLA Popelier with a Contribution from RGA Bone. University of Manchester, 1998.

56 Matta, C. F. *J. Phys. Chem. A* **105** (2001), 11088.

57 Koritsánszky, T., Volkov, A., and Coppens, P. *Acta Cryst. A* **58** (2002), 464.

58 Dittrich, B., Hübschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D., and Luger, P. *Acta Cryst. A* **61** (2005), 314.

59 Coppens, P., Sabine, T. M., Delaplane, R. G., and Ibers, J. A. *Acta Cryst. B* **25** (1969), 2451.

60 Frey, M. N., Lehmann, M. S., Koetzle, T. F., and Hamilton, W. C. *Acta Cryst. B* **29** (1973), 876.

61 Agranat, I., Caner, H., and Caldwell, J. *Nature Rev. Drug Discov.* **1** (2002), 753.

62 Flack, H. D. *Acta Cryst. A* **39** (1983), 876.

63 Betteridge, P. W., Carruthers, J. R., Cooper, R. I., Prout, K., and Watkin, D. J. *J. Appl. Cryst.* **36** (2003), 1487.

64 Sheldrick, G. M. SHELXL-97. A Program for Refinement of Crystal Structures. University of Göttingen, 1997.

65 Bernardinelli, G., and Flack, H. D. *Acta Cryst. A* **41** (1985), 500.

66 Dittrich, B., Strümpel, M., Koritsánszky, T., Schäfer, M., and Spackman, M. A. *Acta Cryst. A* **62** (2006), 217.

67 Dittrich, B., Hübschle, C. B., Luger, P., and Spackman, M. A. *Acta Cryst. D* **62** (2006), 1325.

Part IV
Chemical Bonding and Reactivity

13
Interactions Involving Metals: From "Chemical Categories" to QTAIM, and Backwards

Piero Macchi and Angelo Sironi

13.1
Introduction

The Lewis model [1] enables description of bonds between main group elements but its extension to the realm of transition metal compounds is difficult because many subtle effects complicate solution of the chemical bonding puzzle. Interpretation is equally uncertain for weak intermolecular interactions. Similarly, the quantum theory of atoms in molecules (QTAIM) [2] readily visualizes the intuitive concepts of atoms and bonds whereever the Lewis theory works whereas for transition metal complexes the topology of the electron density $\rho(\mathbf{r})$ reveals chemical bonding features in a less straightforward manner.

More effort is therefore necessary to build up the "correspondence rules" linking QTAIM to traditional bonding concepts. As soon as a proper model is devised, however, significant information can be retrieved, resulting in new understanding or even some re-thought of "classical" chemical bonds involving transition metals. This is the objective of this chapter. Some examples from the recent literature will be discussed. To enable unbiased comparison we reproduce all theoretical calculations at the same level of theory, B3LYP/6-311++G(2d,2p), except for metal hydride bridges, for which it has been demonstrated [3] that B3LYP [4] functional is less reliable than BP86 [5].

13.2
The Electron Density in Isolated Metal Atoms – Hints of Anomalies

What is the electron density of an isolated atom telling us about that atom? Many atomic properties can be predicted from a knowledge of the distribution of the ground-state electron density of the isolated element. For example, the concentration of the valence electrons, and their separation from the core, can be used to predict bond distances, the tendency to form covalent bonding, the degree of localization of the bonding electrons, etc.

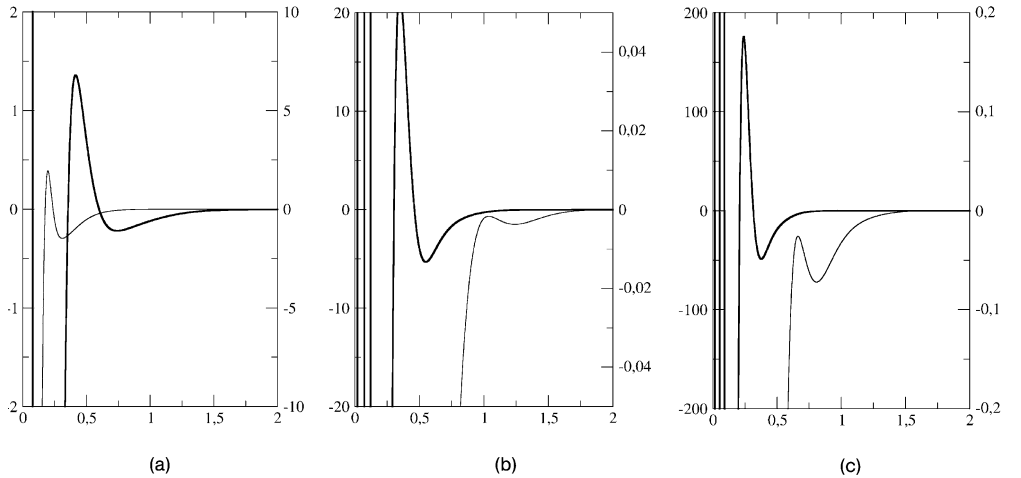


Fig. 13.1 H(**r**) (solid line) and L(**r**) (bold line) radial profiles of the isolated N (a), Mn (b) and As (c) atoms. **r** is in Å, H(**r**) and L(**r**) are in a.u. (left scale for L, right scale for H). L(**r**) of As has a ripple maximum of its outermost shell at ca. $r = 1.3$ Å (this is missing in Mn).

As pointed out by Bader [2, 6], the negative Laplacian ($\mathrm{L}(\mathbf{r}) = -\nabla^2\rho(\mathbf{r})$) of an isolated ground-state atomic density [7] reproduces the electronic shell structure, alternating positive and negative regions with a maximum and a minimum for each shell [8]. However, heavy atoms deviate from this behavior [9]: the expected maxima and minima of the N shell are not observed for the elements from Sc to Ge, but from As to Kr the M and N shells are again separated, although the outermost maxima do not necessarily have $\mathrm{L}(\mathbf{r}) > 0$ (the radial L(**r**) for N, Mn, and As are shown in Fig. 13.1). Similar trends are observed for the elements of successive rows and have been justified on the basis of the diffuse character of the outermost electrons [10b]. The electron density $\rho(\mathbf{r})$ decreases regularly to zero, as also do the kinetic energy density G(**r**) and the potential energy density V(**r**) in absolute values (being everywhere negative). The total energy density H(**r**) ($= \mathrm{V}(\mathbf{r}) + \mathrm{G}(\mathbf{r})$) is not monotonic, however, although it does not have the same structure as L(**r**).

Because of to the local virial theorem [2]:

$$(\hbar^2/4m)\nabla^2\rho(\mathbf{r}) = 2G(\mathbf{r}) + V(\mathbf{r}) \tag{1}$$

H(**r**) can be rewritten as:

$$H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r}) = 1/2((\hbar^2/4m)\nabla^2\rho(\mathbf{r}) + V(\mathbf{r})) \tag{2}$$

showing it is somewhat reminiscent of the features of the Laplacian distribution: an atom usually has an inner (core) region and an outer (valence) shell, irre-

spective of its principal quantum number. The boarders might occur at $H(\mathbf{r}) > 0$ and the "atomic" limit at large $\mathbf{r}$ (Fig. 13.1).

When a chemical bond is formed, all these functions are no longer spherical and their properties characterize atomic interactions. In particular, $L(\mathbf{r})$ enhances the features of the charge distribution and electron pair localization, providing a physical connection with the classical Lewis model and the valence shell electron-pair repulsion theory [11]. $L(\mathbf{r})$, however, remains affected by the distribution in the unperturbed atom. For example, when a covalent bond links two second-row atoms, each has a distinct *valence shell charge concentration* (VSCC, maxima of $L(\mathbf{r})$) along the bond path, quite close to the location of the corresponding $L(\mathbf{r})$ valence shell maxima in the isolated atomic density. Because the outermost shells of the two bonded atoms partially overlap, we have $L(\mathbf{r}) > 0$ in the bond region, typical of an *open-shell* interaction. Naively, this enables visualization of the *electron sharing mechanism*; this is, however, better explained physically by means of the bond delocalization index (δ) [12, 13] related to the electron pairs shared between two atoms (Table 13.1). Along the bond path, $V(\mathbf{r})$ dominates $G(\mathbf{r})$, hence $H(\mathbf{r}) < 0$. The regions where the electrons are localized are also visible in the $H(\mathbf{r})$ distribution, see for example Fig. 13.2a, in which the $H(\mathbf{r})$ distribution of C_2H_6 is

Table 13.1 Properties of the electron-density distribution in some reference molecules discussed in the text.

Compound	Bond, A–B	d_{A-B} (Å)[a]	$d_{A\text{-}bcp}$ (Å)[a]	$d_{B\text{-}bcp}$ (Å)[a]	ρ_b (e Å^{-3})	$\nabla^2\rho_b$ (e Å^{-5})	H_b/ρ_b (h e^{-1})	G_b/ρ_b (h e^{-1})	$\oint_{A\cap B}\rho(r)$ (e Å^{-1})[b]	δ(A, B) [SCF][c]
H_2	H–H	0.743	0.371	0.371	1.822	−29.2	−1.12	0.00	1.42	1.00
C_2H_6	C–C	1.529	0.765	0.765	1.597	−11.9	−0.75	0.23	2.15	1.00
C_2H_4	C=C	1.326	0.663	0.663	2.381	−27.1	−1.19	0.39	2.98	1.89
NaF	Na–F	1.942	0.902	1.040	0.342	9.92	0.27	1.76	0.53	0.27
Ne_2	Ne–Ne	3.795	1.898	1.898	0.002	0.06	0.86	1.18	0.004	0.002
CO	C≡O	1.126	0.386	0.740	3.416	5.31	−1.92	2.03	3.23	1.81
H_3BCO	B–C	1.521	0.495	1.025	1.004	10.42	−0.85	1.58	2.06	0.47
H_3BCO	C≡O	1.130	0.390	0.740	3.366	4.87	−1.89	1.99	3.21	1.64
$Cr(CO)_6$	Cr–C	1.927	0.952	0.976	0.717	11.2	−0.28	1.37	2.12	0.83
$Cr(CO)_6$	C≡O	1.140	0.392	0.748	3.291	3.09	−1.89	1.96	3.23	1.62
$Co_2(CO)_6(NH_3)_2$	Co–N	2.063	0.973	1.090	0.502	8.40	−0.14	1.31	1.42	0.51
$Co_2(CO)_6(PH_3)_2$	Co–P	2.190	1.013	1.177	0.628	4.10	−0.34	0.80	1.98	0.81
$Co_2(CO)_6(AsH_3)_2$	Co–As	2.320	1.044	1.277	0.512	3.52	−0.30	0.78	1.76	0.71
Na_2	Na–Na	3.052	1.526	1.526	0.063	−0.03	−0.13	0.10	0.52	1.00
$Co_2(CO)_6(AsH_3)_2$	Co–Co	2.693	1.346	1.346	0.244	0.10	−0.30	0.33	1.58	0.52

[a] For each bond, d_{A-B}, $d_{A\text{-}bcp}$, and $d_{B\text{-}bcp}$ are the A–B interatomic distances and the distances of atoms A and B, respectively, from the bcp.
[b] $\oint_{A\cap B}\rho(\mathbf{r})$ is the electron density integrated over the (A|B) interatomic surface.
[c] δ(A, B) is the delocalization index between atoms A and B.

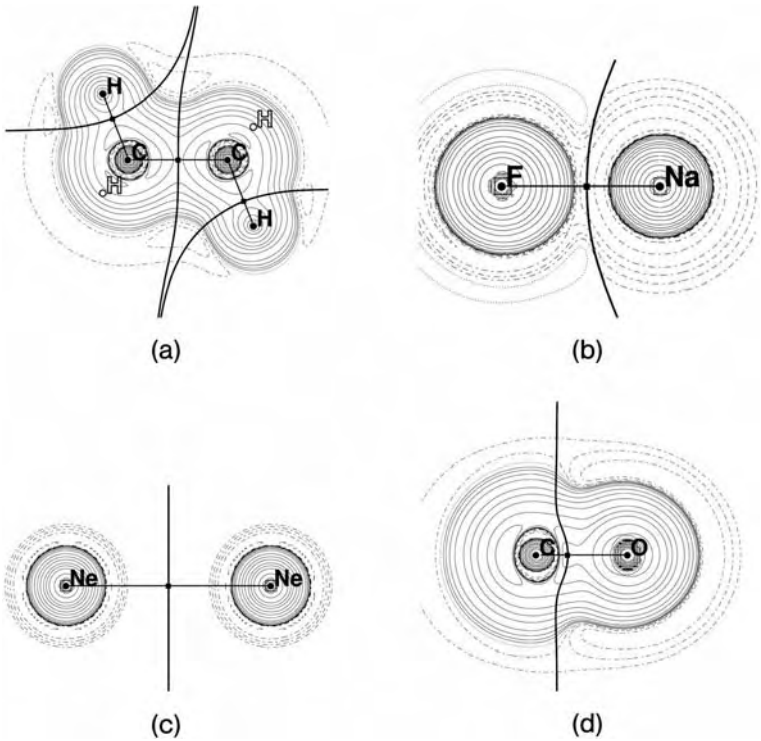


Fig. 13.2 H(**r**) distributions in (a) C_2H_6, (b) NaF, (c) Ne_2, (d) CO (contours increase in value in the order $\pm x \times 10^n$ a.u., where $x = 1.0, 2.0, 4.0, 8.0$ and $-3 \leq n \leq +2$; solid lines are negative contours). Bold labels and dots represent atoms lying in the plane; empty symbols represent atoms projected on to the plot plane. IAS and molecular graphs are superimposed.

shown. The topology of L(**r**) defines the so called *atomic graph* around each atom in a molecule, which typically reproduces its hybridization state [2].

If, in contrast, an ionic bond is created, as in NaF, the outermost shells of the cation and the anion are superimposed. The *bcp* is shifted towards Na in a large and flat negative region of L(**r**) where $\rho(\mathbf{r})$ is small (Table 13.1). The electronic configurations of F and Na now coincide with that of Ne and are better approximated by two *closed shells*. Only a small electron delocalization is actually observed at the *bcp* and G(**r**) dominates in the valence region; thus a large area of $H(\mathbf{r}) > 0$ separates the two atoms (Fig. 13.2b).

Although the correlation between L(**r**) and covalent interactions seems valid, Cremer and Kraka [14] were the first to note it could not suffice for weaker covalent bonds and, instead, they proposed investigating whether the electron density in the bonding region stabilized the system, for example by analyzing H(**r**) at the *bcp* or over the whole interatomic surface. The implicit assumption is that molec-

ular density would be defined by the boundary of $H(\mathbf{r})$, an idea more recently applied by Tsirelson [15], who reconstructed the energy-density plots in ionic and molecular crystals from experimentally derived electron density, by using DFT-like approximations to compute $G(\mathbf{r})$ [15, 16], hence $H(\mathbf{r})$, by use of Eq. (1). The plots in Fig. 13.2 demonstrate that the intuition of Cremer and Kraka was probably correct. This approach can be even more useful when applied to transition metal chemistry.

Bader also considered the energy-density quantities in detail, suggesting that $H(\mathbf{r})$, $G(\mathbf{r})$, and $V(\mathbf{r})$ be measured on the basis of the total amount of electron density, with which they clearly correlate. Indicators such as H_b/ρ_b or G_b/ρ_b thus inform us whether the potential or kinetic energy density is dominating at the *bcp* or whether $G(\mathbf{r})$ is in local excess with respect to the electron density itself. G_r/ρ_r usually increases in correspondence with a "shell closure" – "external" if two closed-shell atoms are interacting or "internal" at short interatomic distances, for example in multiple bonds, in which the two cores become close (this is discussed in more detail elsewhere [17]).

When the bond is covalent but is significantly polarized the interatomic surface is shifted toward the VSCC of the least electronegative atom and, eventually, both VSCCs belong to the basin of the most electronegative atom. In CO the *bcp* is so close to C that $\nabla^2\rho_b > 0$ and $G_b/\rho_b > 0$. Nevertheless, the $H(\mathbf{r})$ is everywhere negative (Fig. 13.2). The shift of the *bcp* significantly affects the delocalization index between C and O, which no longer represents the expected number of Lewis bonded pairs (Table 13.1), counting of which in the presence of charge transfer is not possible solely on the basis of the pair density [13].

13.3 Two-center Bonding

As introduced in the previous section, the chemistry of main-group elements is usually characterized by ionic or covalent bonds. The latter are typically discussed in terms of interactions between hybridized atomic orbitals, in which each binding partner contributes one electron to the bonding pair. This picture is substantially valid for metal–metal bonds also, but according to the most popular theories on transition metal chemistry (for example the ligand-field theory, LFT) [18], different genealogy is required when describing two-center (*2c*) metal–ligand bonding [19] resulting from interaction of doubly occupied orbitals on the ligand and empty orbitals on the metal (*donation*) or vice versa (*back-donation*).

A different viewpoint was proposed by Bersuker [20], who assigned the nature of two-center–two-electron (*2c–2e*) bonds without resorting to the actual parenthood of the electrons involved in the bonding but differentiating between valence (localized) and coordination (delocalized) bonds. This scheme is, apparently, closer to the actual phenomenon of the electron density because, in principle, when a chemical bond is formed the participating electrons lose their original affiliation. "However, the genealogic concepts seems to be quite convoluted also

within the framework of QTAIM, because the bonding to transition metals is characterised by a "low overlap" [18] and thus quite reminiscent of the electron density distribution at atomic level."

Following the genealogy scheme, we analyze two different *2c* bonds to a metal: the classical coordination of a ligand (i.e. the donation of one electron pair) and metal–metal bonding (i.e. pairing of two electrons).

13.3.1 The Dative Bond

A classical donor–acceptor bond involving main-group atoms is, for example, $OC–BH_3$ [21]. The Laplacian distribution, the bond paths, and the interatomic surfaces are shown in Fig. 13.3a. A unique VSCC, corresponding to the lone pair of the nucleophilic atom (C) and charge depletion at the electrophilic atom (B) are found along the bond path. The interatomic surface is shifted toward the acceptor and lies inside the charge-depletion shell of B. Therefore, $\nabla^2\rho_{\mathbf{b}} > 0$, although significant deformation has occurred relative to the constituting molecular frag-

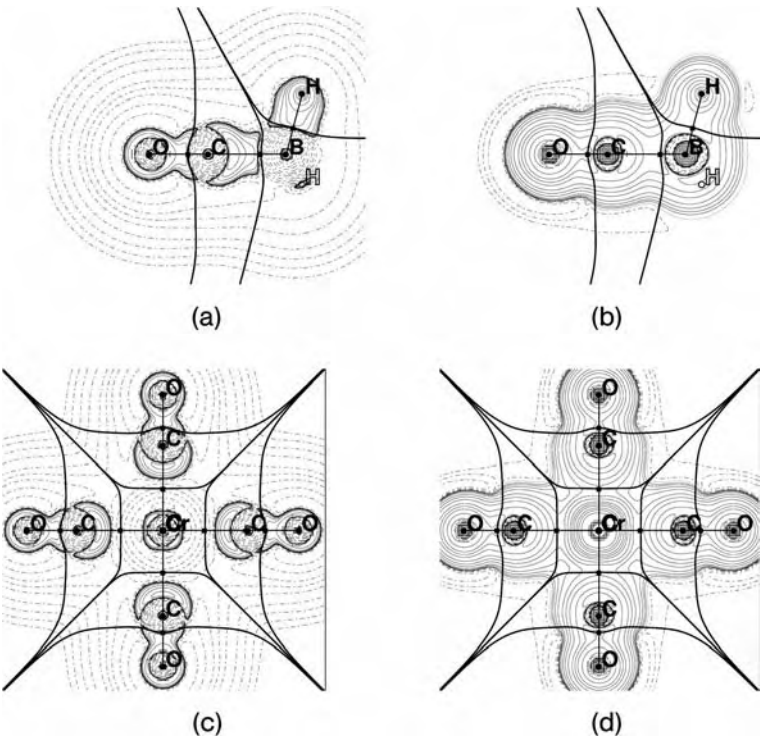


Fig. 13.3 $\nabla^2\rho(\mathbf{r})$ and $H(\mathbf{r})$ distributions in H_3BCO (a and b) and $(CO)_5Cr(CO)$ (c and d). Contours and labeling are as in Fig. 13.2.

ments. The presence of some polarization undermines a simple electrostatic model for this bond. In fact, although $G(\mathbf{r})$ is certainly large, it does not overwhelm $V(\mathbf{r})$, hence $H(\mathbf{r})$ has a negative continuum between B and C (Fig. 13.3b). As anticipated above, interactions in which polarity plays an important role are characterized by significantly reduced delocalization indexes, $\delta(B, C) = 0.47$.

From a different perspective, some energy partitioning schemes [22] predicted that the electrostatic energy affords the largest stabilization in these complexes. It is conceivable that the electrostatic and the covalent terms are closely connected and mutually reinforced, and thus more difficult to separate by analyzing $\rho(\mathbf{r})$. $H(\mathbf{r})$, however, cannot inform us which effect dominates but rather on the occurrence of some stabilization as a result of orbital interaction [14].

No inverse effect (i.e. from B to C) is observable, because the C–O bond indicators (including the distance) are not affected by the coordination and delocalization between O and B is very small ($\delta = 0.03$). There is, however, some electron sharing between the C and the H atoms ($\delta = 0.24$), because of the delocalized nature of the bonding inside the BH_3 unit.

13.3.1.1 Metal Carbonyls

The binary transition-metal carbonyl complexes $M(CO)_n$ are the most studied organometallic molecules, and are often used as reference materials to describe the bonding of transition metals in low oxidation states. The bonding in $Cr(CO)_6$, $Fe(CO)_5$, and $Ni(CO)_4$ is usually understood in terms of the Dewar–Chatt–Duncanson mechanism [23], which invokes reciprocal donation of electron pairs (one from the C lone pair and one from correctly oriented metal *d* orbitals). The electrons used in the bonding are quite localized in nonbonded regions of the isolated unperturbed fragments, i.e. $M(CO)_{n-1}$ and CO. Because of the above-mentioned genealogy, this localization is only marginally affected by the formation of a bond (Fig. 13.3c). Electron delocalization between the metal and the carbonyl is, in fact, approximately 1.0 (Table 13.1), even though more than one electron pair is formally involved in the bonding.

The M–C interatomic surface lies in a region of charge depletion even more pronounced than that in $OC–BH_3$, although associated with $H(\mathbf{r}) < 0$ (Fig. 13.3d). Again the dichotomy between covalent and electrostatic interaction is mixed, without easy deconvolution [24].

The three complexes introduced above are also representative of three different types of stereochemistry (octahedral, bipyramidal, and tetrahedral) and electronic configurations (d^6, d^8 and d^{10}) of zerovalent metals. In LFT terms they correspond to different splitting of *d*-orbitals (t_{2g}/e_g; $e''/e'/a_1$; e_g/t_{2g}) [18], which affects $L(\mathbf{r})$ distribution of the isolated metals in their corresponding electronic configuration [10]. The atomic graphs of the metals predicted at theoretical level and those derived from accurate experimental electron density determination were compared in a recent study [25]. In $Cr(CO)_6$ the atomic graph of Cr is cubic in shape [eight vertexes, twelve edges, and six faces]. The eight vertexes are $(3, -3)$ critical points of $L(\mathbf{r})$ (i.e. charge accumulations produced by the t_{2g} orbitals) and the six faces are centered by $(3, +3)$ critical points (i.e. charge depletion

because of the empty e_g orbitals, disposed in the direction of the incoming ligand density). This configuration has been associated with a "lock and key" mechanism [26], but is basically the visual representation of the LFT prediction. The atomic graphs of $Fe(CO)_5$ and $Ni(CO)_4$ have, respectively, trigonal prismatic [6, 9, 5] and octahedral [6, 12, 8] shapes, though somewhat dependent on the density model employed [25]. Atomic charges show that back-donation is in the order $Cr > Fe \approx Ni$, in agreement with computed and observed stretching frequencies.

Cortés-Guzmán and Bader [27] related the atomic quadrupole moments of C and O to the amount of σ-donation and π-back-donation. The changes occurring to quadrupole moments of a carbonyl axially coordinated to a metal are expected to address the accumulation of density along the axis (because of σ-donation) in contrast with that perpendicular to the axis, occupying a torus around the bound atom (indicative of increased π-density on the carbonyl C). Although the theoretical quadrupole moments were confirmed by the experimental determination [25], some caution is necessary because the atomic volumes of C atoms change quite substantially on coordination.

13.3.1.2 Donor–Acceptor Interactions of Heavy Elements

Many complexes contain a donor–acceptor interaction between a heavy main-group element and a metal. One might expect $L(\mathbf{r})$ to be affected by the small amount of charge concentrated in the atomic valence shell of the donor, and questions about the real nature of such interactions could arise. For this reason we studied the As–Co bond in $Co_2(CO)_6(AsPh_3)_2$ [28]. Although the region of negative Laplacian around As is very small in theoretical maps, and absent from experimental maps, we found many similarities with more classical donor–acceptor bonds. If the whole class of $Co_2(CO)_6(XH_3)_2$ molecules (X = N, P, As) is considered (Fig. 13.4 and Table 13.1) we note they share very similar features of $H(\mathbf{r})$ distribution and that the Laplacian lobe corresponding to the location of the donor electron pair on X decreases for the heavy elements. Despite this, the topological indexes do not depict the Co–As bond as a weak interaction: the delocalization is larger than in N–Co and only slightly smaller than in P–Co. We can, moreover, appreciate the tight correlation between the H_b/ρ_b and δ(Co, X).

13.3.2 Direct Metal–Metal Bonding

The most studied metal–metal bonds are those in homoleptic $M_2(CO)_n$ dimers or in some homo or heteroleptic small metal clusters. The presence of direct chemical bonding between two metals was the subject of discussion for many years. When dealing with "unsupported" M–M contacts, lack of bridging ligands and the 18-electron rule drive chemists to speak of *direct* covalent (as in $Mn_2(CO)_{10}$) or even dative (as in $CrOs(CO)_{10}$) [29] M–M bonds. Some doubts arose, however, when semiempirical calculations were performed to address 1,3 M···CO interactions as the major source of attraction [30]. Previous accurate electron-density determinations by X-ray diffraction [31, 32] were unable to furnish a solution to the

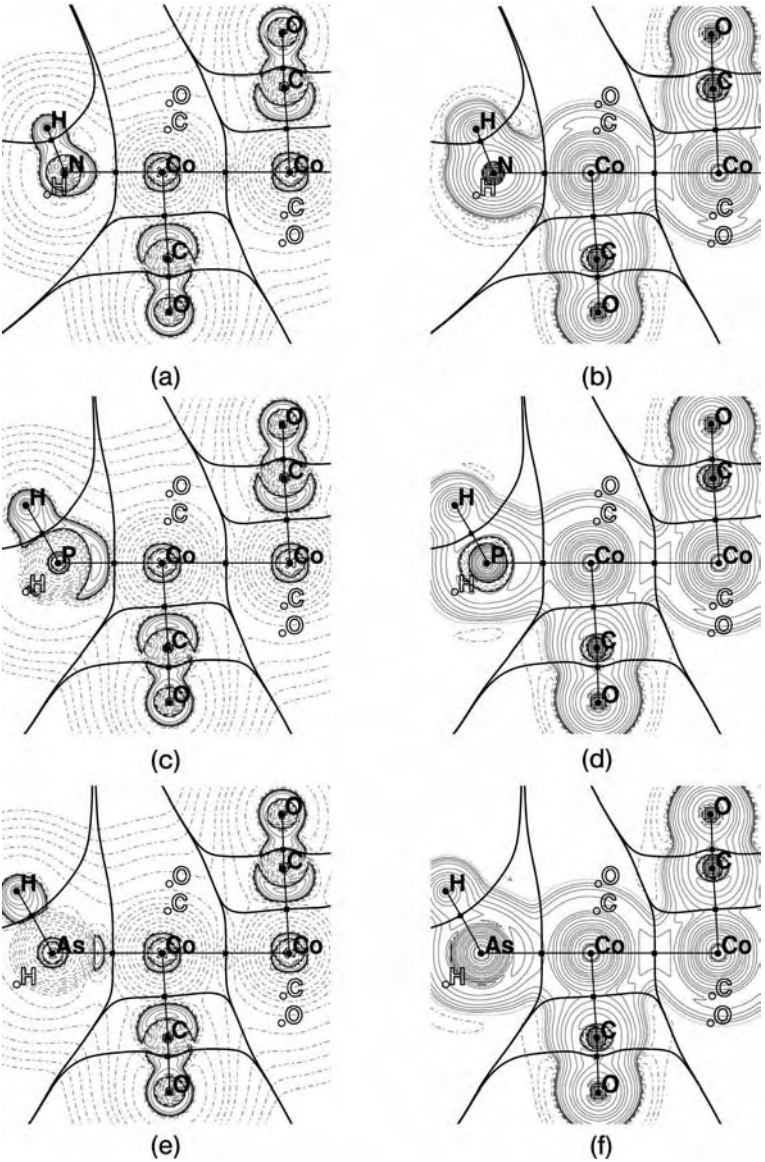


Fig. 13.4 $\nabla^2\rho(\mathbf{r})$ and $H(\mathbf{r})$ distributions in $[Co_2(CO)_6(NH_3)_2]$ (a and b), $[Co_2(CO)_6(PH_3)_2]$ (c and d), and $[Co_2(CO)_6(AsH_3)_2]$ (e and f). Contours and labeling are as in Fig. 13.2.

puzzle. In fact, interpretation of rather noisy deformation density maps could not produce a clear picture of the M–M interactions. Theoretical deformation maps, also, could not reveal bonding effects, because density accumulation in M–M bonds is visible only through *fragment deformation maps* [33], i.e. using as promolecule the superimposition of computed $[ML_n]^\bullet$ fragments rather than spherical atoms.

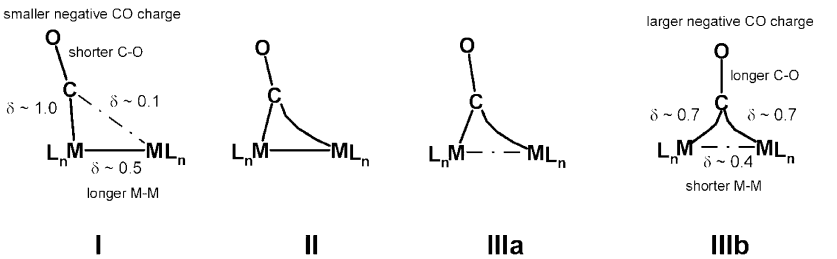
QTAIM later offered, instead, a less ambiguous understanding of the M–M interactions, with clear distinction between unsupported species (in which an M–M bond path is found) and ligand bridged species (in which a M–M bond path is usually *not* found) [10, 34]. QTAIM was used to interpret the experimental charge density of $Mn_2(CO)_{10}$ [35, 36] and $Co_2(CO)_6(AsPh_3)_2$ [28] that confirmed the presence of a bond path linking the two metals. While Macchi et al. [28] and Farrugia et al. [36] considered the M–M interactions as genuine covalent bonds (based especially on Cremer and Kraka's criterion), however, [14] Bianchi and coworkers classified the Mn–Mn bond as *metallic* with features "between ionic and covalent" [35b]. The main argument was the positive Laplacian found at the M–M *bcp*. On the basis of a similar reasoning, Uhl et al. classified the Ni–Ni interaction in $CpNi(\mu\text{-}InCH_3)_2NiCp$ as closed-shell [37].

Inspection of the Laplacian distribution along an M–M bond path shows that the *bcp* is located in an $L(\mathbf{r})$ maximum (although $L_b < 0$) produced by condensation of the two vanishing N shells (if first period transition metals are concerned) [10]. This is somewhat similar to what occurs in simple metal or semimetal diatomic molecules (for example Na_2 or B_2) [17]. The potential energy density still dominates at the *bcp* (thus $H_b < 0$) and the total amount of kinetic energy density per electron is small ($G_b/\rho_b \ll 1$). Despite the small ρ_b, the M–M interaction is not necessarily weak; the density integrated over the whole zero-flux surface separating two bonded atoms provides meaningful results when diffuse electrons contribute to the bond (e.g. Na_2 in Table 13.1). This concept was stressed in the analysis of $Co_2(CO)_6(AsPh_3)_2$ [28], in which classification problems based on "traditional" QTAIM rules were first addressed. Since then, many other bonding indicators and classification schemes have been proposed, sometimes leading to controversial interpretation. For example, Gervasio et al. [38] applied, to transition metals, the scheme of Espinosa and coworkers [39], who classified bonding interactions according to the spatial region they occupy: the shortest interatomic separation is characteristic of pure shared-shell bonds ($\nabla^2\rho_b < 0$ and $H_b < 0$), the longest is characteristic of pure closed-shell bonds ($\nabla^2\rho_b > 0$ and $H_b > 0$), and the central region ($\nabla^2\rho_b > 0$ and $H_b < 0$) is called a *transit* region. All interactions with transition metals fall in the intermediate region [38], as one might expect on the basis of their properties at the isolated atom level. This of course suggests M–M and M–L bonds cannot be pure closed shell bonds, as originally proposed [35a], and that these interactions are weaker than covalent bonds between main group atoms. Gervasio et al. [38] also used the *flatness* criterion, $f = \rho_{min}/(\rho_{bcp})_{max}$, [40] to classify chemical bonds in solids: a large flatness was associated with closed-shell bonds and a low flatness with stronger covalent bonds. This criterion cannot be applied to isolated molecules, because ρ_{min} would

simply be zero. This limitation has been tackled [38] by analysis of a supported M–M bond system, $Co_2(CO)_6(\mu\text{-}CO)(\mu\text{-}C_4H_2O_2)$, with an unexpected M–M bond and a ring-type molecular graph. ρ_{rcp} was therefore taken as the ρ_{min} reference for the flatness; it therefore seemed to be quite large, as in closed-shell bonds. The bias introduced is quite severe and would not enable study of an unsupported system. Accurate analysis performed by Farrugia [34c] also revealed that no MCM ring structure is actually present in $Co_2(CO)_6(\mu\text{-}CO)(\mu\text{-}C_4H_2O_2)$, because the direct Co–Co bond path is missing; this invalidated the conclusions of Ref. [38].

It is worth remarking that the terms "closed-shell character" or "open-shell character" of a chemical bond refer to the electronic configuration of the "parent" fragments. When the bond is formed, such rigorous partition is no longer possible and the electron density itself cannot provide this information, except by resorting to a phenomenological correspondence with prototype situations. It was in this way that Bader [2] derived, for second and third-row molecules, a simple "translation" of the orbital concepts using $\nabla^2\rho(\mathbf{r})$ (Table 13.1). However, the absence of VSCC for some heavier elements undermines this scheme [41]. The concept of "shared interaction" mainly implies *the concerted movement of a certain fraction of electrons in two (or more) atomic basins*. Within QTAIM, the correct indicator for such a concept is the delocalization index δ, whose definition is based on the presence of the same electron pair(s) in two (or more) atomic basins. The delocalization index is defined on the basis of pair-density distribution and is, therefore, not accessible from experimental X-ray diffraction experiments, in contrast with the electron density [42]. Correspondence with electron density defined properties is, therefore, necessary for interpretation of results from experimental work.

In Table 13.1 it is apparent that the metal–metal bond in unsupported $(XH_3)(CO)_3Co\text{–}Co(CO)_3(XH_3)$ molecules is associated with $\delta \approx 0.5$, which would violate the assumption that the two metals share one electron pair to satisfy the 18-electron rule. We noted, however, that the electron-sharing process in metal carbonyl dimers is rather complex, because it involves not only the two metals but also the equatorial carbonyls (Scheme 13.1(I)) [17]. This is in agreement with the later finding by Ponec et al. [43] on the basis of application of domain-averaged Fermi holes theory to $Mn_2(CO)_{10}$: the Mn–Mn bond has the character



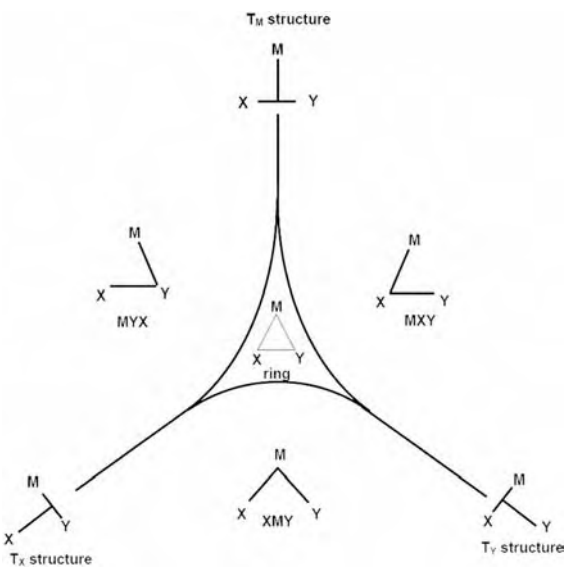
Scheme 13.1 M–M–CO and M–CO–M bonding in transition metal carbonyl clusters.

of an ordinary covalent single σ-bond but evidence of Mn···(CO) intramolecular interactions is also observed. It has been suggested that the shape of the M–C bond paths (slightly bent toward the other metal) and the Laplacian of the equatorial carbonyls may be evidence of a contribution from CO to the M–M [28]. It should be noted that all these observations transfer the problem of the M–M bond in $M_2(CO)_n$ to multicentered bonding. We will consider this more explicitly in Section 13.4.3 on supported metal–metal bonds.

Bonds between alkali metals in Na_2 or K_2, which share with Co–Co of $Co_2(CO)_6(XH_3)_2$ the small ρ_b and the positive $\nabla^2\rho_b$, but have delocalization indexes able to reproduce the expected bonding electron pair involved, might be regarded as truly unsupported M–M bonds. Once again, the concept of covalence is apparent from H_b (< 0) and from the small amount of kinetic energy density per electron ($G_b/\rho_b \leq 0.5$ he^{-1}).

13.4
Three-center Bonding

The interactions among three atoms is associated with a variety of molecular graphs (Scheme 13.2) [44] that might represent different modes of bonding, although this commonly accepted interpretation should be treated with caution.



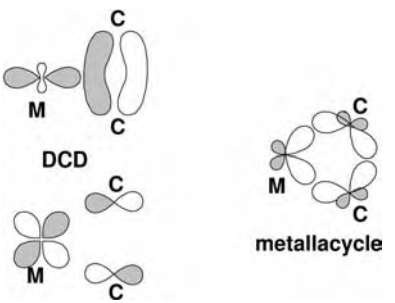
Scheme 13.2 Molecular graphs associated with three-center bonding to a metal atom.

The same molecular graph is, in fact, sometimes associated with different coordination modes (recognized on the basis of other indicators) or different molecular graphs might represent a very similar bonding situation in which minor geometry changes result in different bond path connectivity. In three-center (*3c*) systems, the ring structure is the reference, but not necessarily associated with three separated and localized bonds. Inward curvature of the paths is often an indication of a *3c* delocalized bond, at variance with the ring structure, with outward paths typical of three localized bonds. Depending on the kind of interaction, one can observe ring rupture or the degeneration of bcp and rcp into a T-shape structure, a catastrophe point in the conformational space of XMY systems [2].

In organometallic chemistry *3c* bonding is quite relevant to interpreting the donation of π and σ-density of a preformed bond. This is at the heart, for example, of metal–olefin coordination, agostic interactions, dihydrogen complexes, bridging hydrides, and carbonyl supported metal–metal interactions, all presented below.

13.4.1
π-Complexes

The coordination of an olefin to a metal is one of the most studied chemical bonds involving transition metals – a prototypical bond with enormous implications, for example in catalysis. Many theoretical investigations have been conducted to understand the features of the M–η^2(C=C) bond, leading to the dichotomy between Dewar–Chatt–Duncanson (*DCD*) donor–acceptor complexes and metallacycles. The total number of electrons involved in the bonding is the same, but their spatial localization (and the hypothetical spin state of the interacting fragments) is quite different. In the *DCD* complex, a bonding electron pair (the π-density of the olefin) is σ-donated to the metal which, in turn, back donates *d*-electrons of π symmetry into the empty π^* of the olefin. In a metallacycle, instead, two localized *2c–2e* M–C bonds are formed at the expense of the C–C double bond character (Scheme 13.3).



Scheme 13.3 MO interaction in DCD and metallacycle coordination.

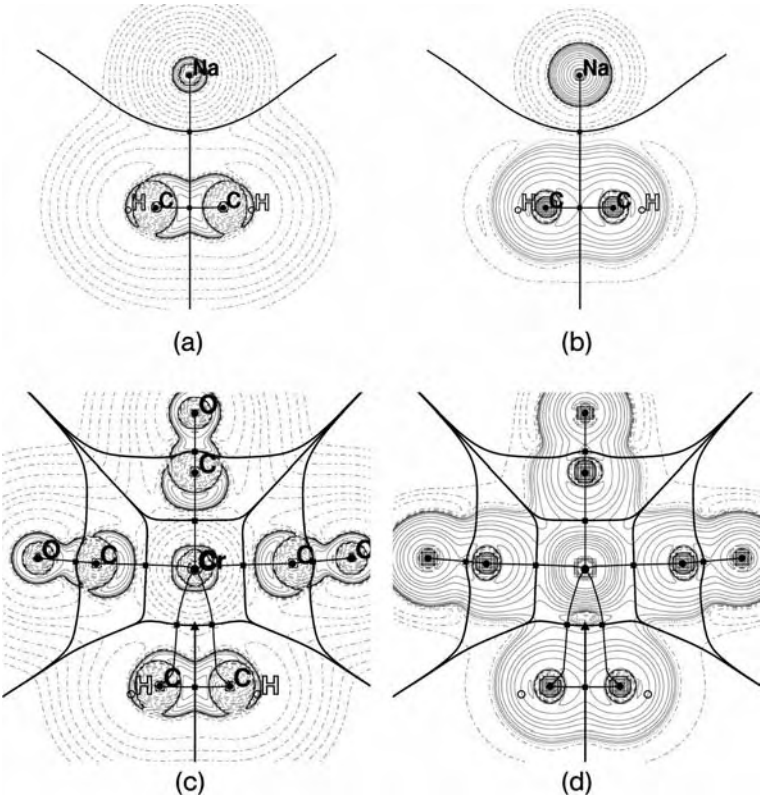


Fig. 13.5 $\nabla^2\rho(\mathbf{r})$ and $H(\mathbf{r})$ distributions in the metal–olefin plane of $Na(C_2H_4)^+$ (a and b) and $Cr(CO)_5(C_2H_4)$ (c and d). Contours and labeling as in Fig. 13.2.

We recognize three main modes of bonding:

1. *Closed shell interaction*: The olefin is coordinated to the metal without significant orbital interaction and binding is achieved by electrostatic interaction between a positively charged metal and the electron density of the C=C bond. One example is olefin coordination to a closed shell cation, as in $Na(C_2H_4)^+$. In Fig. 13.5, we can appreciate the features of the molecular graphs associated with this mode of bonding, characterized by a T-shape graph (i.e. a single line of maximum electron density is connecting Na and the C=C *bcp*). In other words, we are unable to locate separate bonds between the metal and each carbon atom. On the basis of $\nabla^2\rho(\mathbf{r})$ distribution (Fig. 13.5a), the two fragments are almost unchanged from the isolated Na^+ and C_2H_4 and the closed shell nature of the interaction is also revealed by $H(\mathbf{r})$, the negative regions of

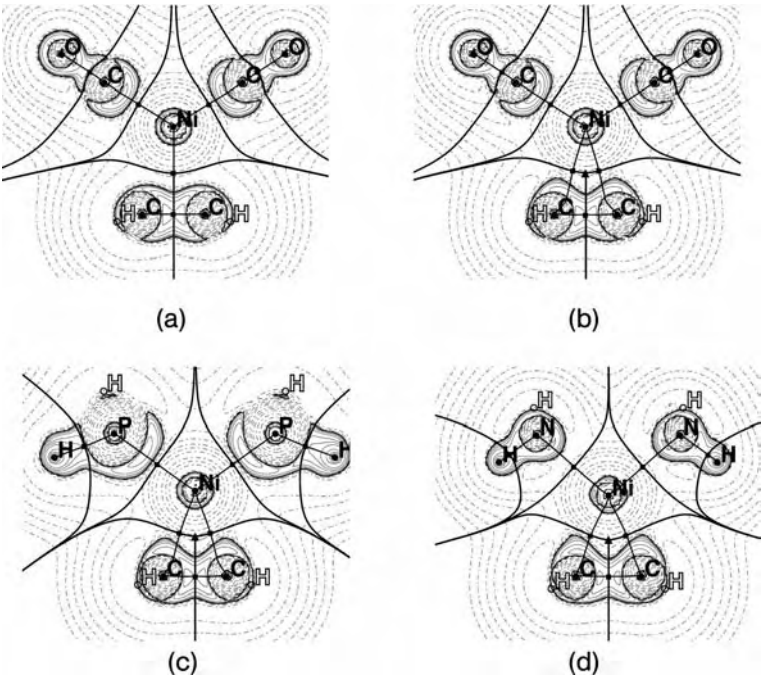


Fig. 13.6 $\nabla^2\rho(\mathbf{r})$ distributions in the metal–olefin plane of: (a) superimposition of $Ni(CO)_2$ and (C_2H_4) densities in the same geometry of the adduct; (b) $Ni(C_2H_4)(CO)_2$; (c) $Ni(C_2H_4)(PH_3)_2$; (d) $Ni(C_2H_4)(NH_3)_2$. Contours and labeling as in Fig. 13.2.

which inside the two fragments are separated by a large area of $H(\mathbf{r}) > 0$ (Fig. 13.5b). Despite a very low covalence, was shown by Frenking et al. [45] that this mode of coordination is associated with a large stabilization energy (relative to the constituting fragments).

2. *Dewar Chatt Duncanson (DCD) ring complex*: If back-donation and donation are active but relatively small, the π electron density is still primarily involved in the C=C bond but the p orbitals are slightly rehybridized and some electron density gradients of the C atoms join the interatomic surface of Ni forming two separate inwardly curved or "*endo*cyclic" M–C bond paths. The more abundant the donation, the straighter the bond paths will appear (Fig. 13.6). This can be emphasized by drawing the contribution to $\nabla\rho(\mathbf{r})$ of the π-orbital of the free olefin as reported in Fig. 13.7. In free ethylene, at short range around each C atom $|\nabla\rho(\mathbf{r})|$ has a mirror-symmetric four-lobe shape. If the unperturbed olefin density is superimposed to the ML_n density, a T-shape

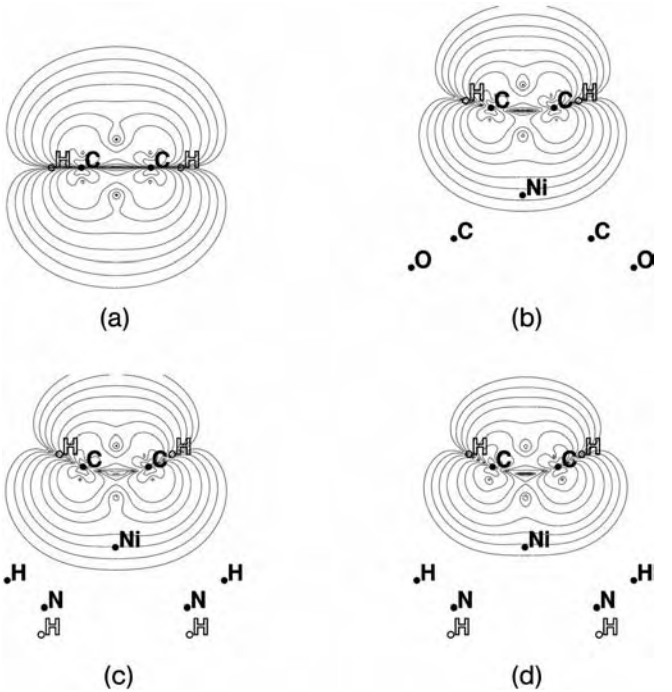


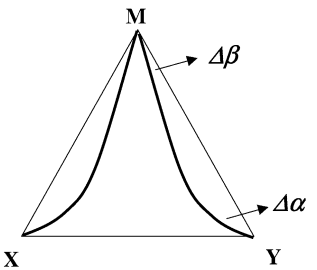
Fig. 13.7 Side view of $|\nabla\rho(\mathbf{r})|$ generated by the π density in C_2H_4 in different electronic states and geometries: (a) ground state (^1A_g) in an isolated molecule; (b) $^1A_1$, as in $(CO)_2Ni(C_2H_4)$; (c) $^1A_1$, as in $(NH_3)_2Ni(C_2H_4)$; (d) $^3B_2$ as in $(NH_3)_2Ni(C_2H_4)$.

molecular graph appears, but when the *DCD* mechanism is switched on, the mirror symmetry of $|\nabla\rho(\mathbf{r})|$ is lost, the $|\nabla\rho(\mathbf{r})|$ lobes undergo a "disrotation" (Fig. 13.7b), and the molecular graph assumes a ring shape (Fig. 13.6).

3. *The metallacycle*: On increasing the metal–olefin interaction, two localized M–C bonds are formed and the C=C π-density is broken. The benchmark molecular graph is cyclopropane, characterized by *exocyclic* bond paths [46], although the different electronegativity and stereochemical requirements of C and M do not enable this limit to be reached. The topology of metallacycles is substantially similar to that depicted in point 2, but the charge concentrations along the M–C bonds are more pronounced and the concavities of the bond paths at C are reduced. The *DCD* model no longer works, because the perturbation is now too large and a complete spin reallocation has produced two M–C covalent interactions. Indeed, a proper ring shape structure is retrieved from the superposition of C_2H_4 density in the

triplet state ($^3B_2$) on to the ML_n density. The olefin HOMO orbital is associated with more *exocyclical* gradients (Fig. 13.7c) and the bond order of M–C and C–C bonds becomes much closer in agreement with the delocalization indexes.

Based on QTAIM and charge decomposition analysis, Frenking [47] concluded that the *DCD* ring complex structure is quite typical of metals in lower oxidation states. It is apparent from Table 13.2, however, that for some 16 or 18-electron complexes of zero-valent metals the range of C–C distances span from shorter (<1.40, hence closer to the classical *DCD* complex) to larger (>1.44). The main rationale seems to be the acidity of the metal fragment: if the metal is coordinated to electron-withdrawing ligands (for example acidic CO) metal back-donation is poor and the olefin coordination is weaker; in contrast, electron-donating groups opposed to the olefin induce larger back-donation. Accordingly we see that C–C–M bond path angles approach the geometrical angles (see



Scheme 13.4 Definition of bond path differences in MX_2 systems.

Scheme 13.4 for definitions).

The QTAIM data for a series of metal–olefin complexes emphasize the tight correlation between C=C elongation (hence C rehybridization) and the tendency to produce more separated M–C bond paths (Table 13.2). This was previously formulated in experimental determination of $\rho(\mathbf{r})$ in a typical *DCD* complex, ($Ni(COD)_2$) [48] (COD = 1,5-cyclooctadiene). Scherer and coworkers [49], on the other hand, have recently drawn, for some 16-electron $Ni(XR_3)_2(C_2H_4)$ complexes (X = N, P; R = alkyl, H), the density (and Laplacian) of the molecular orbital mainly responsible of the olefin-to metal-donation. Their outward pointing shapes would suggest that *endocyclic* bond paths are not the correct markers of donation. Donation of the paired π-density (point 2 above) is, however, always mixed with that of the unpaired sp^3-like density (point 3) thus any molecular orbital cannot be representative of the pure *DCD* bonding mode, especially because the examples chosen are ahead in the metallacycle conversion (Table 13.2). The reasoning in point 2 is based on the expected behavior of the olefin not contaminated by its triplet configuration and it is valid if the objective of molecular graph analysis is to retrieve a given electronic configuration out of a multiconfigurational system.

Table 13.2 Topological analysis for three-center MX_2 systems representative of π and σ-donation. $\Delta\alpha$ and $\Delta\beta$ angles refer to Scheme 13.4.

Compound	X–X					M–X					X	M
	d (Å)	ρ_b (e Å^{-3})	$\nabla^2\rho_b$ (e Å^{-5})	H_b/ρ_b (h e^{-1})	δ (X, X)	d (Å)	ρ_b (e Å^{-3})	$\nabla^2\rho_b$ (e Å^{-5})	H_b/ρ_b (h e^{-1})	δ (M, X)	$\Delta\alpha$ (°)	$\Delta\beta$ (°)
$Na(C_2H_4)^+$	1.336	2.342	−26.3	−1.16	1.85	2.729	0.098	1.34	0.14	∼0[e]	–	–
$(CO)_5Cr(C_2H_4)$	1.365	2.201	−23.2	−1.09	1.52	2.400	0.293	3.03	−0.11	0.29	26.9	−27.8
$(CO)_2Ni(C_2H_4)$	1.385	2.103	−20.7	−1.04	1.42	2.05	0.599	5.48	−0.30	0.56	22.5	11.2
$(CO)_4Fe(C_2H_4)(\|\|)$	1.399	2.053	−19.9	−1.01	1.33	2.151	0.510	4.51	−0.24	0.52	20.5	13.3
$(CO)_4Fe(C_2H_4)(\perp)$[d]	1.399	2.054	−19.8	−1.01	1.38	2.151	0.499	4.93	−0.23	0.43	32.5	2.9
$(PH_3)_2Ni(C_2H_4)$[a]	1.400	2.034	−19.2	−1.00	1.35	2.01	0.65	5.77	−0.31	0.64	10.4	9.5
$(NH_3)_2Ni(C_2H_4)$[b]	1.438	1.873	−15.8	−0.92	1.22	1.925	0.781	6.29	−0.35	0.80	2.1	17.4
$(NH_3)_2Ni + (C_2H_4)$[c]	1.438	1.912	−15.7	−0.91	–	1.925	0.727	8.38	−0.18	–	–	–
$(CO)_2(NH_3)_2Fe(C_2H_4)$	1.451	1.841	−15.5	−0.89	1.16	2.070	0.603	4.95	−0.28	0.68	0.3	24.6
$(CO)_5Cr(H_2)$	0.799	1.603	−22.0	−1.01	0.70	1.796	0.424	8.70	−0.02	0.26	–	–
$(PH_3)_5Cr(H_2)$	0.858	1.392	−16.0	−0.918	0.61	1.710	0.512	8.36	−0.16	0.40	49.1	−26.1
$(PH_3)_5Cr(H_2)^\dagger$	1.565	0.352	1.00	−0.154	0.27	1.831	0.447	2.82	−0.22	0.55	20.5	−21.5
$(PH_3)_5Cr(H)(H)$	–	–	–	–	–	1.641	0.658	3.91	−0.37	0.64	–	–

[a] The stable conformation is actually of C_s symmetry (but C_{2v} is here reported for comparison with the other structures).
[b] The stable conformation is actually of C_2 symmetry (but C_{2v} is here reported for comparison with the other structures).
[c] Superimposition of $Ni(NH_3)_2$ and C_2H_4 (ground states) in the same geometry of the $Ni(NH_3)_2(C_2H_4)$ optimized complex.
[d] The molecular geometry of the $\perp$ conformation was not optimized because no stationary point can be located and it converts easily into the || conformation. Thus the || geometry was kept fixed and the C_2H_4 simply rotated about the twofold axis.
[e] The definition of the Na basin was not obtainable and integration was performed only at a very approximate level.

In the same work, Scherer et al. discussed other interesting features of the bond paths at the metal atom. In a classical *DCD* complex the acceptor metal d-orbital is empty and usually bisects the C–M–C angle, whereas the orbitals involved in back-donation are *exocyclically* directed, for example $Cr(CO)_5(C_2H_4)$ in Fig. 13.5 and Table 13.2. Thus, one should expect outward curvature at the metal, as one can appreciate from the $\nabla^2\rho(\mathbf{r})$ distribution, which is indicative of charge accumulation by the Cr 3d shell to lie outside the $Cr–C_2$ ring. This is not observed for late transition metals, however, particularly the d^{10} elements [28, 49], which have all d orbitals occupied and therefore should have a spherical Laplacian. The back-donation process induces an aspherical density, with charge depletion corresponding with the orbitals involved. Thus, in contrast with Cr, in Ni complexes there is a dominant negative region of the Laplacian associated with an orbital *not* involved in the back-bonding (and thus completely filled) and pointing toward the center of the MC_2 ring. Consequently, the bond paths at Ni (or even at Fe) are slightly *endocyclic*. Scherer and coworkers also noted that around the midpoints of the Ni–C bonds the path curvatures change again (becoming *exocyclic*), because there the π-back-bonding orbitals dominate. It should be noted that this is not the rule, but an exception for electron-richer metal.

The Laplacian usually shows two distinct charge concentrations, corresponding with the M–C bonds, which become larger as the M–C bond strengthens, in agreement with the shape of the bond paths and with the larger amount of electron density at bcp. Accordingly, M–C and C=C delocalization indexes reproduce quite well the reallocation of the olefin π-density in M–C interactions. Eventually, in coordination mode 3, the distribution of total energy density is characterized by a unique negative region incorporating both the metal and the olefin fragment. All the arguments proposed above can be used to identify the different bonding contributions in parallel ($\parallel$) or perpendicular ($\perp$) coordination of an olefin in the equatorial plane of $Fe(CO)_4(C_2H_4)$. From simple orbital reasoning [50] the parallel stereochemistry is favored, because back-donation is more active. Assuming the same coordination geometries for parallel and perpendicular $Fe(CO)_4(C_2H_4)$, we can appreciate from Table 13.2 that the parallel conformer has the features associated with larger back-donation and stronger M–C interactions (i.e. smaller δ(C=C), larger δ(M–C), and smaller inward curvature of the bond paths).

13.4.2
σ-Complexes

Although the donation of π density to a metal was discovered and studied much earlier than that of σ density, π and σ-complexes share many similarities. As far as hydrogen is concerned, σ-complexes span dihydrogen compounds, agostic interactions, and bridging hydrides [51], the bonding of which is usually described in terms of 3c–2e bonds. There is now general consensus that 3c–2e systems may have many different geometries, take some 3c–4e character (depending from the relevance of π-donation from the metal), and, eventually, evolve into two *2c–2e* bonds.

13.4.2.1 Dihydrogen and Dihydride Coordination

The first dihydrogen complexes ($W(CO)_3(PR_3)_2(H_2)$, R = cyclohexyl, isopropyl) were reported in 1984 by Kubas et al. [52]. Some pioneering QTAIM studies on η^2 coordination of H_2 to a metal were published in the early nineties [53, 54]. In the dihydrogen complexes the mechanism of coordination is reminiscent of the *DCD* model, although based on donation of σ-density. Accordingly, whereas the metal–olefin coordination is characterized by the π-complex–metallacycle dichotomy, the dihydrogen complexes have similar uncertainty with regard to dissociation into dihydride. Similarly, there are weakly bound complexes, observed when the metal center is quite acidic and characterized by a $\mathbf{T_M}$ molecular graph (e.g. $Cr(CO)_5(H_2)$), and strongly bound complexes, characterized by more substantial back-donation and a genuine ring graph (e.g. $Cr(PH_3)_5(H_2)$). Because the H atoms lack of directional orbitals, however, the M–H bond paths of the ring structure are quite inwardly curved and the H_2 density is only weakly polarized in the direction of the metal. Dapprich and Frenking [53] reported a ring structure (though almost collapsed) for $(CO)_5M(H_2)$ molecules (M = Cr, Mo, W) based on MP2 and CCSD(T) calculations with effective core potential basis sets. The overall picture is, nevertheless, not very different because in both circumstances the dominant interaction is the H–H bond, as demonstrated by the topological and delocalization indexes (Table 13.2), whereas the metal–ligand interaction is just above a pure closed-shell limit (especially in $Cr(CO)_5(H_2)$). On the basis of charge decomposition analysis, Dapprich and Frenking [53] estimated the amount of donation and back-donation, concluding that H_2 is a much weaker π-acceptor than CO, thus justifying the molecular graphs observed.

Tomás et al. [55] showed that a dihydrogen complex can further proceed in the oxidative addition of H_2, which was associated with a small barrier. The final product is an heptacoordinated bipyramidal complex with two hydride ligands in the pentagonal plane, but not in the *cis* configuration. The transition state is a structure containing an H_2 moiety with H···H separation of more than 1.5 Å. In some complexes this structure is actually a stable isomer, because the minimum on the potential energy surface is attained at larger H···H separation, whereas in others an equilibrium between dihydrogen and dihydride can be established [55]. The main chemical problem is then associated with the presence of some residual H···H bonding.

13.4.2.2 Agostic Interactions

Within the QTAIM formalism, Popelier and Logothetis were the first to address the differences between agostic interactions (in d^0 Ti^{IV} complexes) and the 3c–4e hydrogen bonds [56]. They concluded that an agostic interaction is characterized by a bond path with values of the electron density and the Laplacian at the bcp characteristic of an ionic, closed-shell interaction. They proposed criteria based on the variations of atomic H_β properties which, however, lack generality (Table 13.3). Scherer and McGrady [57], in contrast, showed that β-agostic interactions in d^0 metals occur because of delocalization of the M–C_α bond over the metal–alkyl moiety (negative hyperconjugation) rather than the presence of a

Table 13.3 Topological analysis for β-agostic systems and C_2H_6. Variations of Atomic charge (ΔQ), energy (ΔE), volume (ΔV), and dipole (ΔM) of the H_β atom are given relative to those for the reference C_2H_6 molecule. d_A and d_B are distances between the bcp and atoms A and B, respectively. The dhpe complexes are computed in C_s symmetry. The *exo* conformer is not a minimum on the PES of $Cl_3TiCH_2CH_3$ but is here reported for comparative analysis with the *endo* conformer.

β-Agostic system	H_β properties		Bond, A–B	d (Å)	d_A (Å)	d_B (Å)	ρ_b (e Å^{-3})	$\nabla^2\rho_b$ (e Å^{-5})	G_b/ρ_b (h e^{-1})	H_b/ρ_b (h e^{-1})	δ
(dhpe)$Cl_3TiCH_2CH_3$	ΔQ (e)	0.00	Ti–C_α	2.142	1.107	1.038	0.668	0.74	0.46	−0.38	0.70
	ΔE (h.)	−0.00153	C_α–C_β	1.518	0.749	0.769	1.605	−11.5	0.26	−1.23	1.04
	ΔV (Å^3)	−1.78	Ti–H_β	2.143	–	–	–	–	–	–	0.09 (0.10)[a]
	ΔM (D)	−0.074	C_β–H_β	1.118	0.698	0.420	1.749	−20.3	0.19	−1.00	0.86
$[(dhpe)NiCH_2CH_3]^+$	ΔQ (e)	−0.03	Ni–C_α	1.923	0.984	0.941	0.795	3.96	0.73	−0.39	0.82
	ΔE (h.)	−0.00444	C_α–C_β	1.478	0.727	0.751	1.763	−14.3	0.28	−0.85	1.09
	ΔV (Å^3)	−1.44	Ni–H_β	1.676	0.705	0.991	0.533	6.93	1.13	−0.23	0.32 (0.25)[a]
	ΔM (D)	−0.098	C_β–H_β	1.187	0.723	0.465	1.429	−12.8	0.24	−0.86	0.74
$Cl_3TiCH_2CH_3$ exo	ΔQ (e)	+0.02	Ti–C_α	2.040	0.972	1.068	0.856	−0.24	0.44	−0.46	0.84
	ΔE (h.)	+0.00524	C_α–C_β	1.536	0.775	0.761	1.545	−10.9	0.24	−0.73	1.02
	ΔV (Å^3)	−0.30	Ti–H_β	2.875	–	–	–	–	–	–	0.02 (0.05)[a]
	ΔM (D)	−0.021	C_β–H_β	1.093	0.689	0.404	1.888	−24.1	0.16	−1.05	0.94
$Cl_3TiCH_2CH_3$ endo	ΔQ (e)	+0.03	Ti–C_α	2.046	1.070	0.976	0.849	−0.25	0.44	−0.46	0.84
	ΔE (h.)	+0.01187	C_α–C_β	1.527	0.772	0.755	1.574	−11.4	0.24	−0.75	1.02
	ΔV (Å^3)	−0.26	Ti–H_β	4.010	–	–	–	–	–	–	0.02 (0.04)[a]
	ΔM (D)	−0.011	C_β–H_β	1.093	0.690	0.403	1.881	−23.9	0.16	−1.05	0.95
C_2H_6	Q (e)	−0.03	–								
	E (h.)	−0.63765	C_α–C_β	1.529	0.765	0.765	1.597	−11.9	0.23	−0.75	1.00
	V (Å^3)	7.70	–								
	M (D)	0.376	C_β–H_β	1.091	0.685	0.405	1.897	−24.3	0.16	−1.06	0.96

[a] δ(Ti–C_β) in parentheses.

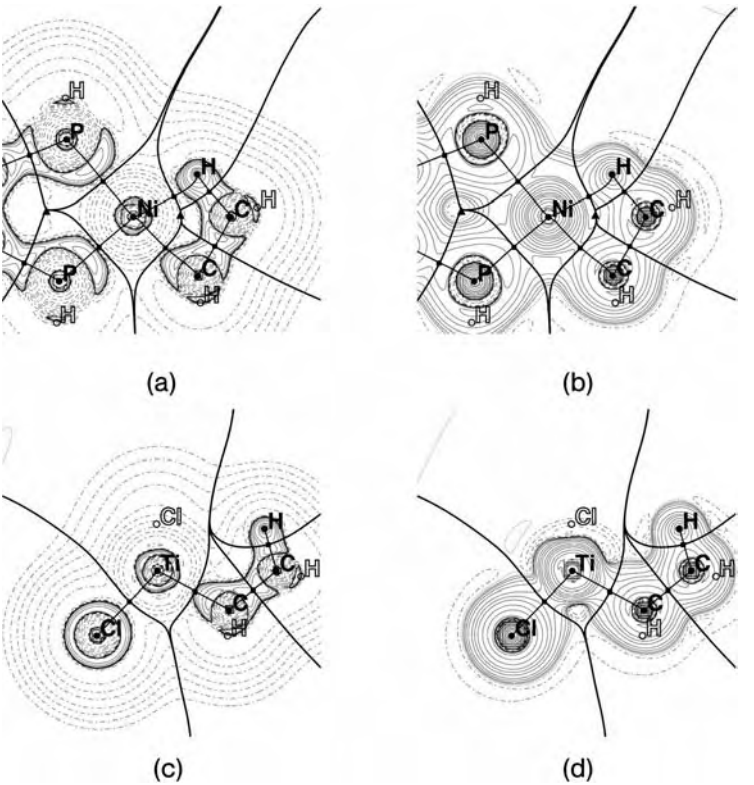


Fig. 13.8 $\nabla^2\rho(\mathbf{r})$ and $H(\mathbf{r})$ distributions in $Ni(dhpe)(C_2H_5)^+$ (a and b) and $TiCl_3(C_2H_5)$ *exo* (c and d).

3c–2e $M\cdots H_\beta–C_\beta$ interaction. In a combined experimental and theoretical study on $Ti(dhpe)Cl_3(C_2H_5)$ ($dhpe = H_2PCH_2CH_2PH_2$) a significantly pronounced bond path for the $Ti–H_\beta$ was not found [58]. In support of this theory, the authors addressed the characteristic curvature of the $Ti–C_\alpha$ bond path (Fig. 13.8) and its ellipticity along the whole path. We also note that delocalization indexes are very informative: if an agostic interaction is activated, both $\delta(M–C_\alpha)$ and $\delta(C_\beta–H_\beta)$ decrease in favor of $\delta(C_\alpha–C_\beta)$ and $\delta(M–H_\beta)$.

It is worth noting that if the main mechanism is hyperconjugative, *exo* or *endo* H_β should both be affected by the agostic interaction. Indeed, the *endo* and *exo* (which is not a minimum on the PES) conformers of $TiCl_3(C_2H_5)$ are characterized by the same features of a weak agostic type interaction. If a $Ti–H_\beta$ bond path is found in the *exo* conformer it is simply because of an additional and, perhaps, negligible source of bonding (i.e. a closed-shell $Ti\cdots H_\beta$ interaction) which does not significantly account for the overall stability of that conformer. The $H(\mathbf{r})$ distribution is in agreement with this view (Fig. 13.8).

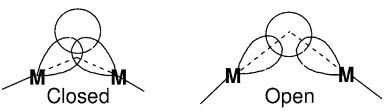
Scherer and McGrady also pointed out that the availability of local Lewis acidity sites on the metal is a feature that favors agostic interactions, as revealed by the VSCCs about the metal atom.

In contrast, electron-richer metals afford real 3c–2e agostic interactions – see the comparison between $(dhpe)Cl_3Ti(CH_2H_5)$ and $[(dhpe)Ni(CH_2H_5)]^+$. As suggested for the $Cp(CO)_2Mn/HSiCl_3$ adduct, however [44, 59], an agostic interaction may transform into (or consist of) the more classical oxidative addition product. This process is of paramount interest in catalysis.

On the basis of these observations we can locate two extremes: the *weak agostic* and *incipient oxidation* adducts. The former is associated with an MYX graph (Scheme 13.2, X = C, Y = H) close to the breaking of the M–Y path, the latter with a ring structure close to the breaking of X–Y to afford an XMY graph. They are also distinguishable on the basis of δ(M–H)/δ(X–H) which is close to zero in the agostic interaction of early transition metals but could reach unity in "symmetric agostic" interactions (Section 13.4.2.3).

13.4.2.3 **Hydride Bridges**

An apparently different type of σ-coordination to a metal is that of hydrides bridging two metal atoms. An experimental and theoretical study of the electron density distribution of $[Cr_2(\mu_2\text{-H})(CO)_{10}]^-$ has shown, however [60], that many similarities with agostic interactions can be drawn for M–H–M systems which, traditionally, were classified on the basis of *open* or *closed* 3c bonding (Scheme 13.5). Stereochemical considerations led to the conclusion that a *closed* system is more adherent to reality, because the direction opposite to the axial ligands point toward the MHM ring center rather than to the hydride, suggesting the presence of some direct M–M bonding. It should be noted that this hypothesis was formulated *before* the agostic interaction was actually discovered.



Scheme 13.5 Hypothesis for bonding in M–H–M systems.

The potential energy surface of $[Cr_2(\mu_2\text{-H})(CO)_{10}]^-$ is particularly flat [3]: the gas phase molecule has C_s symmetry, a nearly symmetric bridge and staggered carbonyl groups (Fig. 13.9) but in the solid state many other conformers can be found as a function of the perturbation induced by the environment; this enables definition of an "experimental" interconversion path between the C_s staggered conformer and a pseudo-D_{4h} (C_{2v}) conformer with eclipsed carbonyl groups and an almost linear M–H–M system (Fig. 13.9). A Cr–Cr bond path is not present and the electron sharing between the two atoms is small and almost constant as a function of Cr–H–Cr bending. The linear geometry, in which no direct metal–

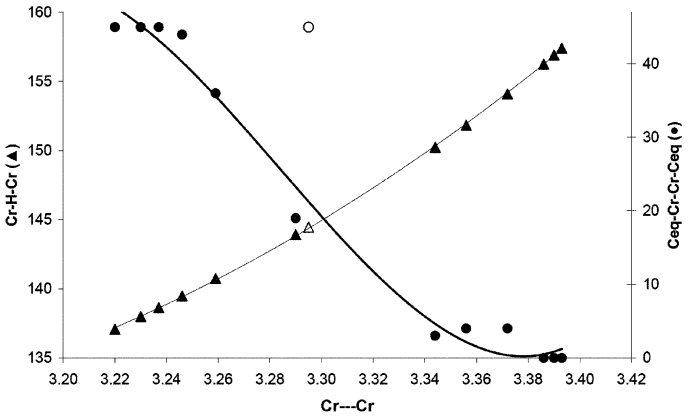


Fig. 13.9 Scatter plot of the dependence of Cr–H–Cr angles (triangles) and C_{eq}–Cr–Cr–C_{eq} torsion (circles) on Cr···Cr distance, taken from the many crystal structures of $[Cr_2(\mu_2\text{-H})(CO)_{10}]^-$ salts. Empty symbols indicate outlier geometry, the nature of which is discussed elsewhere [60].

metal bonding is possible and which contains almost the same amount of delocalization as in the bent structures, should also be noted. QTAIM analysis did not reveal a "unique" effect responsible for stabilization of one isomer over the others. One is tempted to conclude that "packing forces" have a significant effect in determining the solid-state conformation of this anion but some general features emerge. For example, the shift of the H atom out of the idealized $(OC)_{ax}$–M direction seems to be a consistent in M–H–M systems and cannot be attributed to the increased direct M···M interaction. Indeed, an unexpected role played by the nearby equatorial carbonyls has been recognized, because of the features of the Fermi hole density distribution and the corresponding H···CO delocalization indexes (δ approx. 0.1 [60]).

Accurate comparison of $[(CO)_5Cr–H–X]^-$ systems (Table 13.4) revealed much similarity between two X-moieties $Cr(CO)_5$ and BH_3, classically regarded as isolobal [61]. On reducing the acidity of X the limit of a weak agostic interaction can be reached, for example with $X = [CH_3]^+$. Within this framework it is reasonable that, although asymmetric (M···H–C) and symmetric hydride bridge (M–H–M) have undisputedly different geometries, they share a common bonding nature (at least as far as σ-donation of C–H occurs). In more asymmetric (*classical*) agostic interactions, there is a kind of schizophrenic behavior of the metal, the binding of which could be directed alternately toward H or X (or even be absent, as shown above). As the M–H and X–H bonds become more similar in strength, the M–H overwhelms M···X and the structure is characterized by a small M···X delocalization. In the limiting case, $X = Cr(CO)_5$, $\delta(M–H)/\delta(X–H)$ approaches unity and we can speak of a *symmetry-stabilized* agostic interaction [60].

Table 13.4 Summary of QTAIM analysis of X–H and X···Y interactions in X–H–Y and Y–H systems.

X–H–Y	d (Å)	$d_{Y\text{-}bcp}$ (Å)	$d_{bcp\text{-}H}$ (Å)	Y–H–X (°)	Y···X (Å)	$\rho(r)$ (e Å^{-3})	$\nabla^2\rho(r)$ (e Å^{-5})	$G(r)/\rho(r)$ (h e^{-1})	$H(r)/\rho(r)$ (h e^{-1})	$\oint_{A\cap B}\rho$ (e Å^{-1})	ε	δ_{X-H}	δ_{X-Y}	q(H) (e)
$[(CO)_5\mathbf{Cr}–\mathbf{H}]^-$	1.650	1.028	0.622	–	–	0.626	3.58	0.754	−0.354	1.595	0.06	0.59	–	−0.25
D_{4h} $[(CO)_5\mathbf{Cr}–\mathbf{H}–Cr(CO)_5]^-$	1.767	1.063	0.704	180	3.534	0.415	4.24	0.901	−0.184	1.410	0.0	0.38	0.07	−0.37
C_s $[(CO)_5\mathbf{Cr}–\mathbf{H}–Cr(CO)_5]^-$	1.769	1.072	0.698	140.3	3.326	0.437	3.89	0.834	−0.209	1.590	0.02	0.38	0.09	−0.31
$[(CO)_5\mathbf{Cr}–\mathbf{H}–BH_3]^-$	1.752	1.055	0.697	131.5	2.815	0.445	4.59	0.919	−0.196	1.420	0.04	0.38	0.07	−0.44
$[(CO)_5\mathbf{Cr}–\mathbf{H}–SiH_3]$	1.726	1.045	0.681	112.4	2.776	0.466	5.48	1.027	−0.203	1.70	0.02	0.43	0.13	−0.56
$[(CO)_5\mathbf{Cr}–\mathbf{H}–CH_3]$	2.009	1.190	0.830	120.5	2.761	0.211	2.73	0.955	−0.047	0.891	0.27	0.13	0.11	−0.04
$[H_3\mathbf{B}–\mathbf{H}]^-$	1.244	0.543	0.701	–	–	0.993	−3.48	0.73	−0.976	1.689	0.0	0.58	–	−0.64
$[H_3\mathbf{B}–\mathbf{H}–BH_3]^-$	1.322	0.561	0.761	126.6	2.361	0.699	1.44	0.955	−0.811	1.510	0.20	0.35	0.06	−0.57
$[H_3\mathbf{B}–\mathbf{H}–Cr(CO)_5]^-$	1.330	0.575	0.756	131.5	2.815	0.708	−0.11	0.815	−0.826	1.667	0.29	0.36	0.07	−0.44
$[H_3\mathbf{Si}–\mathbf{H}]$	1.491	0.738	0.753	–	–	0.800	2.28	0.868	−0.669	1.477	0.00	0.59	–	−0.62
$[H_3\mathbf{Si}–\mathbf{H}–Cr(CO)_5]$	1.590	0.783	0.812	112.4	2.776	0.631	0.86	0.44	−0.381	0.909	0.18	0.43	0.13	−0.56
$[H_3\mathbf{C}–\mathbf{H}]$	1.096	0.694	0.403	–	–	1.837	−22.62	0.161	−1.023	1.780	0.0	0.98	–	0.003
$[H_3\mathbf{C}–\mathbf{H}–Cr(CO)_5]$	1.122	0.699	0.423	120.5	2.761	1.681	−18.81	0.190	−0.973	2.04	0.003	0.87	0.11	−0.04

13.4.3
Carbonyl-supported Metal–Metal Interactions

Metal cluster scientists have extensively debated the role of direct M–M interaction in carbonyl-bridged 3c–4e metal–metal bonds. This has led to several papers reporting studies of the two most representative molecules, $Fe_2(CO)_9$ and $Co_2(CO)_8$ [17]. The QTAIM shows that the most relevant difference between bridged and unbridged isomers is, respectively, the absence or presence of a M–M bond path (Scheme 13.1(I) and (IIIb)). The topology of a bridged system was first reported for $Co_2(CO)_8$ by MacDougall [10, 33a]. $Fe_2(CO)_9$ was the topic of a long theoretical debate. Whereas empirical rules predict direct Fe–Fe bonding, many MO calculations (semiempirical or ab initio) concluded there was no interaction between the two metals given the small *d* overlap [62], though VB calculations provided more evidence of an interaction [63]. Mealli and Proserpio [64] found that an Fe–Fe bond is formally present even if the through-bond intermetal repulsion overcomes the attractive through-space Fe–Fe interaction. QTAIM analyses have been reported by MacDougall [10] and by Bo et al. [34], who did not find a direct M–M bond path and concluded that "$Fe_2(CO)_9$ is built up by the bridging carbonyls" [34], in agreement with the earlier suggestion by Summerville and Hoffmann [65]. Bo et al. also noted many features of the μ_2-coordination:

- the larger envelope of the valence shell surrounding the carbon, indicative of more delocalized bonding through the metals;
- a larger electronic population on the carbonyl carbon, as a consequence of the better metal-to-ligand charge transfer in the bi-coordinative mode; and
- the presence of two nonbonded VSCCs on the bridging oxygen, indicative of an incipient rehybridization.

Delocalization of the bond was confirmed by analysis of the Fermi hole density maps [34].

The first experimental validation of the electron-density distribution in μ-CO systems came from analysis of the tetrahedral cluster $Co_4(CO)_{11}(PPh_3)$ [66]. In agreement with all-electron HF calculations on the C_{3v} $Co_4(CO)_{12}$ isomer, no direct Co–Co bond path was found for the three bridged edges (the overall topological features of which resembled those of $Co_2(CO)_8$).

To complete the analysis of the hypothetical conversion path from terminal to symmetrically bridging $M_2(CO)$ systems, analysis of a semibridging conformation was undertaken [67]. This study proved that the terminal-to-bridge metamorphosis of a carbonyl, although accompanied by an abrupt change in the molecular graph, actually lies on a type of continuum, especially if the electron-sharing process is considered. The delocalization indexes of M–M and M–C interactions change along the interconversion path in such a way that the overall sharing index is almost constant. This explains both the carbonyl fluxionality in transition metal clusters and the observed continuity of conformations in known M–CO–M

fragments [17]. Thus, even when the carbonyl is terminally bound to just one of the two metal atoms, an M–CO–M system is characterized by *interplay of direct and indirect M–M and M–C interactions* that results in substantial delocalization through the system. This interplay generates different molecular graphs and hampers formation of any truly localized bond, which explains why the M–M unsupported bond has less delocalization than expected (in favor of 1,3 M···CO interactions). In symmetrically bridged metal dimers the view proposed is in agreement with the many indications of M–M bonding (first and foremost the 18-electron rule) because substantial metal–metal electron delocalization is actually present, even in the absence of a direct bond path.

On the basis of H(**r**) distribution and analysis of the contribution of each molecular orbital, Rehinhold et al. [68] came to the conclusion that some direct M–M bonding is actually visible in supported M–M bonds. We note that this view does not contradict the interplay of interactions introduced above and made more visible by the delocalization indexes.

13.5 Concluding Remarks

In the last few years QTAIM has become the model for interpreting theoretical and experimental electron density distributions. Within this framework, the link between bonding modes and topological properties has been fully achieved for molecules of main group atoms. In contrast, the correspondence rules derived cannot be extended in a straightforward manner to organometallic compounds, because bonds to a transition metal have different and much narrower spectrum of topological indexes. Metals are always characterized by diffuse *ns* valence density lacking concentration of charge in the bonding region, even when electron sharing is important. Thus, classifications based on $\nabla^2\rho(\mathbf{r})$ might be misleading or at least incomplete.

On moving from *2c* to *3c* interactions the situation becomes even more complicated, because there are sudden changes in the molecular graph produced by small perturbations of the relative "weight" of each bonding contribution. When a system is close to a catastrophe point, it is the shape rather than the topology of a given molecular graph that is informative, and several examples have been discussed in which the presence of a bond path was actually misleading because it could not completely explain the molecular geometry observed.

We have shown that δ(A, B) and H(**r**) are the most reasonable indicators of the covalent contribution to the total bonding between two (or more) atoms, not least because they are rather insensitive to abrupt changes of the molecular graphs. They are a bridge between traditional chemical categories (mainly formed on the basis of simple MO schemes) and quantum mechanics of the electron density; much work is, however, still required to derive the appropriate correspondence rules, which must take into account subtler details of the molecular graphs.

References

1 G. N. Lewis, *J. Am. Chem. Soc.* **1916**, *38*, 762–785.
2 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, **1990**, Oxford University Press, Oxford.
3 N. A. Richardson, Y. Xie, R. B. King, H. F. Schaefer *J. Phys. Chem. A*, **2001**, *105*, 11134–11143.
4 (a) A. D. Becke, *J. Chem. Phys.*, **1993**, *98*, 5648–5652. (b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B*, **1988**, *37*, 785–789.
5 (a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100. (b) J. P. Perdew, *Phys. Rev. B*, **1986**, *33*, 8822–8824.
6 (a) R. F. W. Bader, H. Essen, *J. Chem. Phys.*, **1984**, *80*, 1943–1960; (b) R. F. W. Bader, P. J. MacDougall, C. D. H. Lau, *J. Am. Chem. Soc.*, **1984**, *106*, 1594–1605.
7 E. Clementi, C. Roetti, *At. Data Nucl. Data Tables* **1974**, 14, 177–478.
8 (a) H. Schmider, R. P. Sagar, V. H. Smith, *J. Chem. Phys.*, **1991**, *94*, 8627–8629; (b) K. D. Sen, M. Slamet, V. Sahni, *Chem. Phys. Lett.*, **1993**, *205*, 313–316.
9 (a) R. P. Sagar, A. C. T., Ku, V. H. Smith, A. M. Simas, *J. Chem. Phys.*, **1988**, *88*, 4367–4374; (b) Z. Shi, R. J. Boyd, *J. Chem. Phys.*, **1988**, *88*, 4375–4377.
10 (a) P. J. MacDougall, *PhD Thesis*, McMaster University, **1990**; (b) P. J. MacDougall, M. B. Hall, *Trans. Am. Crystallogr. Ass.*, **1990**, *26*, 105–123; (c) C. Bo, J. P. Sarasa, J. M. Poblet, *J Phys. Chem.*, **1993**, *97*, 6362–6366.
11 (a) R. J. Gillespie, I. Hargittai, Allyn and Bacon, Boston, MA, **1991**; (b) R. W. F. Bader, G. L. Heard, *J. Chem. Phys.*; **1999**, *111*, 8789–8797.
12 R. F. W. Bader, M. E. Stephens, *J. Am. Chem. Soc.*, **1975**, *97*, 7391–7399.
13 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A*, **1999**, 103, 304–314.
14 (a) D. Cremer, E. Kraka, *Croat. Chem. Acta*, **1984**, *57*, 1259–1281; (b) D. Cremer, E. Kraka, *Angew. Chem., Int. Engl. Ed.*, **1984**, *23*, 627–628.
15 V. G. Tsirelson, *Acta Crystallogr.* **2002**, *B58*, 632–639.
16 Y. A. Abramov, *Acta Crystallogr.*, **1997**, *A53*, 264–272.
17 P. Macchi, A. Sironi, *Coord. Chem. Rev.*, **2003**, *238/239*, 383–412.
18 B. N. Figgis, *Introduction to Ligand Fields*, **1964**, J. Wiley and Sons, New York.
19 G. Frenking, N. Fröhlich, *Chem. Rev.*, **2000**, *100*, 717–774.
20 I. B. Bersuker, *Electronic structure and properties of transition metal compounds*, **1996**, J. Wiley and Sons, New York.
21 J. Volker, G. Frenking, M. T. Reetz, *J. Am. Chem. Soc.*, **1994**, *116*, 8741–8753.
22 K. Morokuma, *Acc. Chem. Res.*, **1977**, *10*, 294–300.
23 (a) J. S. Dewar, *Bull. Soc. Chim. Fr.* **1951**, *18*, C71. (b) J. Chatt., L. A. Duncanson, *J. Chem. Soc.* **1953**, 2939–2947.
24 G. Frenking, K. Wichmann, N. Fröhlich, C. Loschen, M. Lein, J. Frunzke, V. M. Rayón, *Coord. Chem. Rev.* **2003**, *238–239*, 55–82.
25 L. J. Farrugia, C. Evans, *J. Phys. Chem. A*, **2005**, *109*, 8834–8848.
26 T. S. Koritsanszky, P. Coppens, *Chem Rev.*, **2001**, 101, 1583–1627.
27 F. Cortés-Guzmán, R. F. W. Bader, *Coord. Chem. Rev.*, **2005**, *249*, 633–662.
28 P. Macchi, D. M. Proserpio, A. Sironi, *J. Am. Chem. Soc.*, **1998**, *120*, 13429–13435.
29 H. B. Davis, F. W. B. Einstein, P. J. Glavina, T. Jones, R. K. Pomeroy, P. Rushman, *Organometallics.* **1989**, *8*, 1030–1039.
30 D. A. Brown, W. J. Chambers, N. J. Fitzpatrick, R. M. Rawilson, *J. Chem. Soc. A*, **1971**, 720–725.
31 M. Martin, B. Rees, A. Mitschler, *Acta Crystallogr.* **1982**, *B38*, 6–15.
32 P. C. Leung, P. Coppens, *Acta Crystallogr.* **1983**, *B39*, 535–542.

33 (a) A. A. Low, K. L. Kunze, P. J. MacDougall, M. B. Hall, *Inorg. Chem.*, **1991**, *30*, 1079–1086; (b) M. B. Hall, *Electron Distribution and the Chemical Bond*, P. Coppens, M. B. Hall Editors, Plenum Press, New York, **1982**, p. 205.

34 (a) R. Bianchi, G. Gervasio, D. Marabello, *Helv. Chim. Acta*, **2001**, *84*, 722–734; (b) R. Bianchi, G. Gervasio, D. Marabello, *Acta Crystallogr.* **2001**, *B57*, 638–645; (c) L. J. Farrugia, *Chem. Phys. Lett.*, **2005**, *414*, 122–126.

35 (a) R. Bianchi, G. Gervasio, D. Marabello, *Chem. Commun*. **1998**, 1535–1536; (b) R. Bianchi, G. Gervasio, D. Marabello, *Inorg. Chem.* **2000**, *39*, 2360–2366.

36 L. J. Farrugia, P. R. Mallinson, B. Stewart, *Acta Crystallogr.*, **2003**, *B59*, 234–247.

37 W. Uhl, S. Melle, G. Frenking, M. Hartmann, *Inorg. Chem.*, **2001**, *40*, 750–755.

38 G. Gervasio, R. Bianchi, D. Marabello, *Chem. Phys. Lett.*, **2004**, *387*, 481–484.

39 E. Espinosa, I. Alkorta, J. Elguero, E. Molins, *J. Chem. Phys.*, **2002**, *117*, 5529–5542.

40 P. Mori-Sanchez, A. M. Pendás, V. Luaña, *J. Am. Chem. Soc.* **2002**, *124*, 14721–14723.

41 R. F. W. Bader, C. F. Matta, *Inorg. Chem.*, **2001**, *40*, 5603–5611.

42 However, with the recent experimentally constrained wave function approach (D. Jayatilaka, D. J. Grimwood, *Acta Crystallogr.* **2001**, *A57*, 76–86) the pair density is accessible.

43 R. Ponec, G. Yuzhakov, M. R. Sundberg, *J. Comput. Chem.* **2005**, *26*, 447–454.

44 R. F. W. Bader, C. F. Matta, F. Cortés-Guzmán, *Organometallics*, **2004**, *23*, 6253–6263.

45 M. Böhme, T. Wagener, G. Frenking, *J. Organomet. Chem.* **1996**, *520*, 31–43.

46 D. Cremer, E. Kraka, *J. Am. Chem. Soc.*, **1985**, *107*, 3800–3810.

47 (a) U. Pidun, G. Frenking, *Organometallics* **1995**, *14*, 5325–5336; (b) G. Frenking, U. Pidun, *J. Chem. Soc., Dalton Trans.*, **1997**, 1653–1662. (c) R. Stegmann, A. Neuhaus, G. Frenking, *J. Am. Chem. Soc.*, **1993**, *115*, 11930–11938.

48 P. Macchi, D. M. Proserpio, A. Sironi, *J. Am. Chem. Soc.*, **1998**, *120*, 1447–1455.

49 W. Scherer, G. Eickerling, D. Shorokhov, E. Gullo, G. S. McGrady, P. Sirsch, *New J. Chem.* **2006**, *30*, 309–312.

50 T. A. Albright, R. Hoffmann, J. C. Thibeault, D. L. Thorn, *J. Am. Chem. Soc.*, **1979**, *101*, 3801–3812.

51 R. H. Crabtree, *Angew. Chem. Int. Ed.* **1993**, *32*, 789–805.

52 G. J. Kubas, *Metal Dihydrogen and σ-bond Complexes*, **2001**, Kluwer Academic, New York.

53 F. Maseras, A. Lledós, M. Costas, J. M. Poblet, *Organometallics* **1996**, *15*, 2947–2953.

54 S. Dapprich, G. Frenking, *Angew. Chem.*, **1995**, *34*, 354–357.

55 (a) J. Tomàs, A. Lledós, Y. Jean, *Organometallics* **1998**, *17*, 190–195; (b) J. Tomàs, A. Lledós, Y. Jean, *Organometallics* **1998**, *17*, 4932–4939.

56 P. L. A. Popelier, J. Logothetis, *J. Organomet. Chem.*, **1998**, *555*, 101–111.

57 W. Scherer, G. S. McGrady, *Angew. Chem. Int. Ed.*, **2004**, *43*, 1782–1806 and references therein.

58 W. Scherer, W. Hieringer, M. Spiegler, P. Sirsch, G. S. McGrady, A. J. Downs, A. Haalandd, B. Pedersen, *Chem. Commun.*, **1998**, 2471–2472. The original electron density model characterized by a weak bcp was afterwards revised and a Ti–H_β bond path could not be anymore located. Analogously, on changing the level of theory two different molecular graphs are obtained.

59 D. L. Lichtenberger, *Organometallics* **2003**, *22*, 1599–1602.

60 P. Macchi, D. Donghi, A. Sironi, *J. Am. Chem. Soc.*, **2005**, *127*, 16494–16504.

61 R. Hoffmann, *Angew. Chem., Int. Ed. Engl.*, **1982**, *21*, 711–724.

62 C. W. Bauschlicher, *J. Chem. Phys.*, **1986**, *84*, 872–875.

63 L. Pauling, Z. S. Herman, *J. Chem. Educ.*, **1984**, *61*, 582–587.
64 C. Mealli, D. M. Proserpio, *J. Organomet. Chem.*, **1990**, *386*, 203–208.
65 R. H. Summerville, R. Hoffmann, *J. Am. Chem. Soc.*, **1979**, *101*, 3821–3831.
66 P. Macchi, L. Garlaschelli, S. Martinengo, A. Sironi, *J. Am. Chem. Soc.*, **1999**, *121*, 10428–10429.
67 P. Macchi, L. Garlaschelli, A. Sironi, *J. Am. Chem. Soc.*, **2002**, *124*, 14173–14184.
68 O. Kluge, M. Finger, J. Reinhold, *Inorg. Chem.*, **2005**, *44*, 6494–6496.

14
Applications of the Quantum Theory of Atoms in Molecules in Organic Chemistry – Charge Distribution, Conformational Analysis and Molecular Interactions

Jesús Hernández-Trujillo, Fernando Cortés-Guzmán, and Gabriel Cuevas

14.1
Introduction

This chapter deals with description of the structure and reactivity of several organic molecules by use of the quantum theory of atoms in molecules (QTAIM) [1]. For some topics, emphasis is put on the information that can be obtained from the bond critical points (BCPs) and the delocalization of the Fermi hole density, as accounted for by the delocalization index between two atoms. Relationships between these two types of descriptor are discussed to illustrate how the QTAIM can account for their known experimental chemical behavior. The rest of the chapter analyzes the $^3J_{HH}$ coupling constants of some aliphatic and aromatic molecules (Section 14.2), conformational processes including rotational barriers of XCH_2CH_2X molecules and the anomeric effect of heterocyclohexanes (Section 14.3), and the electronic structure of aromatic molecules (Section 14.4). Finally, closing remarks from a global perspective are presented in Section 14.5.

14.2
Electron Delocalization

14.2.1
The Pair-density

Information about the electronic structure of a molecule in a given state is provided by the wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_N)$ of the N-electron system, were $\mathbf{x} = (\mathbf{r}, s)$, represents the space and spin coordinates of an electron. According to the postulates of quantum mechanics, $\Psi(\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_N)$, contains all the information that can be known about the system [2, 3]. As an alternative, characterization can also be performed in terms of the properties of the electron-density functions derived from the wavefunction. One of these provides the probability density for

finding any of the N electrons in a volume element that includes both its space and spin coordinates, $d\mathbf{x}_1$, irrespective of the position and spin of the remaining electrons:

$$\gamma(\mathbf{x}_1) = N \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\, d\mathbf{x}_2 \ldots d\mathbf{x}_N \tag{1}$$

By means of an integration of this function over the spin coordinate s_1 one obtains the one-density:

$$\rho(\mathbf{r}_1) = \int \gamma(\mathbf{x}_1)\, ds_1 \tag{2}$$

i.e. the probability density of finding an electron in a spatial volume element $d\mathbf{r}_1$. Integration of $\rho(\mathbf{r}_1)$ over all space yields the number of electrons in the system, N. This scalar field is the electron density obtained experimentally from crystallographic experiments. Because the underlying theory of the topological properties of the electron density has been discussed in detail elsewhere [1], the rest of this subsection will be devoted to a brief description of the pair density, because its use for analysis of electron delocalization from the viewpoint of the QTAIM is spread throughout several papers. The pair density is defined by:

$$\pi(\mathbf{x}_1, \mathbf{x}_2) = \frac{N(N-1)}{2} \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\, d\mathbf{x}_3 \ldots d\mathbf{x}_N \tag{3}$$

This function determines the probability of finding any two electrons in space–spin volume elements $d\mathbf{x}_1$ and $d\mathbf{x}_2$, irrespective of the spin and position of the remaining electrons. The corresponding spinless function is obtained when integration over the spin coordinates s_1 and s_2 is performed. Thus:

$$P(\mathbf{r}_1, \mathbf{r}_2) = \int \pi(\mathbf{x}_1, \mathbf{x}_2)\, ds_1\, ds_2 \tag{4}$$

represents the probability of any two electrons being one at $\mathbf{r}_1$ and the other at $\mathbf{r}_2$ simultaneously, irrespective of the position of the remaining electrons and of their spin. Because the molecular Hamiltonian involves only one-electron and two-electron operators, no higher than pair density functions are necessary to describe the interactions. Double integration of $P(\mathbf{r}_1, \mathbf{r}_2)$ yields the number of electron pairs of the N-electron system. From these equations, $\rho(\mathbf{r}_1)$ can be obtained from $P(\mathbf{r}_1, \mathbf{r}_2)$ by integration as follows:

$$\rho(\mathbf{r}_1) = \frac{2}{N-1} \int P(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_2 \tag{5}$$

The one-density function can be expressed in terms of spin contributions:

$$\rho(\mathbf{r}_1) = \rho^{\alpha}(\mathbf{r}_1) + \rho^{\beta}(\mathbf{r}_1) \tag{6}$$

in which the spin state of the electrons is given by the spin functions α and β. The spin density can also be defined as the difference $\rho^{\alpha}(\mathbf{r}_1) - \rho^{\beta}(\mathbf{r}_2)$, a scalar field that accounts for the excess of spin up (α) over spin down (β) contributions to the one-density.

$P(\mathbf{r}_1, \mathbf{r}_2)$ can also be resolved into its spin contributions by integration of $\pi(\mathbf{x}_1, \mathbf{x}_2)$ over all spin possibilities:

$$P(\mathbf{r}_1, \mathbf{r}_2) = P^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2) + P^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) + P^{\beta\alpha}(\mathbf{r}_1, \mathbf{r}_2) + P^{\beta\beta}(\mathbf{r}_1, \mathbf{r}_2) \tag{7}$$

For example, $P^{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2)$ is the probability density of having an electron at $d\mathbf{r}_1$ with spin α and another at $d\mathbf{r}_2$ with spin β. In addition, the conditional probability of finding an electron at $\mathbf{r}_1$ if another electron is at $\mathbf{r}_2$, irrespective of the position of the remaining electrons, is obtained from the pair density:

$$\frac{P_2(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} = \rho(\mathbf{r}_2)[1 + h(\mathbf{r}_1, \mathbf{r}_2)] \tag{8}$$

The term $h(\mathbf{r}_1, \mathbf{r}_2)$ is a distribution function determined by inclusion of Coulomb correlation and, to a greater extent, by the spin distribution of the electrons, as required by the Pauli exclusion principle. In the absence of Coulomb correlation, exchange correlation is the only one present and $h_x(\mathbf{r}_1, \mathbf{r}_2)$ is called the Fermi hole density in which the subindex x emphasizes its exchange correlation dependence. This density vanishes for the contributions of α–β interactions, because the motion of two electrons with different spin is uncorrelated and is different from zero for two electrons with the same spin. At the restricted Hartree–Fock level (RHF), the Fermi hole density for α electrons is [4]:

$$h_x^{\alpha}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{\sum_{i,j} \phi_i^*(\mathbf{r}_1)\phi_j^*(\mathbf{r}_2)\phi_j(\mathbf{r}_1)\phi_i(\mathbf{r}_2)}{\rho^{\alpha}(\mathbf{r}_1)}. \tag{9}$$

Double integration of the product $\rho^{\alpha}(\mathbf{r}_1)h_x{}^{\alpha}(\mathbf{r}_1, \mathbf{r}_2)$ over the basin of a given atom A defined by the QTAIM provides a measure of the number of electrons located on that atom, whereas double integration over the basins of two different atoms A and B in the molecule, not necessarily sharing an interatomic surface, accounts for the number of electrons shared between them. In this manner, localization and delocalization indices $\lambda(A)$ and $\delta(A, B)$, respectively, are defined by [5]:

$$\lambda(A) = \sum_{i,j} \eta_i^{1/2}\eta_j^{1/2}S_{ij}(A)^2, \quad \delta(A, B) = 2\sum_{i,j} \eta_i^{1/2}\eta_j^{1/2}S_{ij}(A)S_{ij}(B) \tag{10}$$

In these definitions, η_i and η_j denote the occupation numbers of natural orbitals ϕ_i and ϕ_j and $S_{ij}(A)$ is the overlap integral of ϕ_i and ϕ_j over the basin of atom A. For RHF, the summations run over the occupied orbitals, $\eta_i = \eta_j = 2$, and reduce to [6]:

$$\lambda(A) = 2\sum_{i,j} S_{ij}(A)^2, \quad \delta(A,B) = 4\sum_{i,j} S_{ij}(A)S_{ij}(B) \tag{11}$$

These latter expressions have also been used in the Kohn–Sham approximation (KS) and it has been found that the corresponding numerical values are similar to those found with RHF [7], although its use can be criticized because of the N-representability problem of KS orbitals.

The localization and delocalization indices add up to the atomic population, $N(A) = \lambda(A) + \frac{1}{2}\sum_{B\neq A}\delta(A,B)$, and the total number of electrons of atom A that are delocalized is $\Delta(\mathrm{A}) = \mathrm{N(A)} - \lambda(\mathrm{A})$. The delocalization of the Fermi hole has been used to explain several chemical effects. Of these, two important examples are:

1. their relationship with the Lewis model of electron pairs [6]; and
2. their ability to provide a physical foundation for the VSERP model of molecular geometry [8].

In addition, the delocalization indices and the properties of the one-density, which in what follows will be referred to as the electron density $\rho(\mathbf{r})$, enable appropriate characterization of the chemical bond, as is illustrated in the examples discussed in this chapter.

14.2.2
$^3J_{HH}$ Coupling Constants and Electron Delocalization

One direct application of the delocalization index is in the study of coupling constants between vicinal H atoms of organic molecules. Beginning from the Hamiltonian for the electron interactions in the field of nuclei with magnetic moments, four main contributions are identified for the nuclear–nuclear spin-coupling constants – the Fermi contact, the paramagnetic spin–orbit, the spin–dipolar, and the diamagnetic spin–orbit interactions [9]. Of these, the Fermi contact contribution has frequently been found to be dominant, for example for protons not directly bonded to each other. For this example, by use of molecular orbital theory [10], it has been found that the coupling constant, J, for nuclei with coordinates $\mathbf{R}_n$ and $\mathbf{R}_{n'}$ is proportional to $\sum_{i,j}\phi_i^*(\mathbf{R}_n)\phi_j^*(\mathbf{R}_{n'})\phi_j(\mathbf{R}_n)\phi_i(\mathbf{R}_{n'})$ and, from Eq. (9), to the product $\rho^{\alpha}(\mathbf{R}_n)\mathrm{h}^{\alpha}(\mathbf{R}_n,\mathbf{R}_{n'})$. J has, moreover, also been written in terms of $P^{\alpha\beta}(\mathbf{R}_n,\mathbf{R}_{n'}) - P^{\alpha\alpha}(\mathbf{R}_n,\mathbf{R}_{n'})$, the excess number of β over α electrons at $\mathbf{R}_{n'}$ given that there is an α electron at $\mathbf{R}_n$; this difference can be related to orbital-based definitions of bond order [10]. Consequently, the Fermi contact contribution to $^3J_{HH}$ results mainly from the coupling of nuclear spins mediated by the electronic

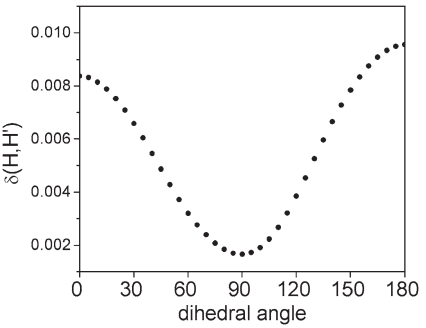


Fig. 14.1 Relationship between δ(H, H′) and the H–C–C–H′ dihedral angle for ethane. QCISD/6-311G(*d*,*p*) wavefunctions used.

spins and, to a great extent, from the values of the *s*-type orbitals at the nuclear positions. From this discussion, it is reasonable to assume proportionality between $^3J_{HH'}$ and δ(H, H′) [11], an assumption that is supported by the empirical correlations described below.

An important feature of the Fermi contact contribution is its sensitivity to geometric changes. This is illustrated by the well-known Karplus-type behavior [12] of $^3J_{HH'}$ of ethane as a function of the H–C–C–H dihedral angle, ϕ – the $^3J_{HH'}$ values obtained from valence bond calculations were successfully subject to nonlinear fitting of type $\cos^2\phi$. Interestingly, the corresponding values of δ(H, H′) shown in Fig. 14.1 have the same trend as a function of ϕ [4]. Accordingly, nonlinear fit of δ(H, H′) against ϕ of the type $\delta(H, H') = A\cos^2\phi + B$ can be achieved. Following Karplus [12], the fitting yields the data given in Table 14.1 for ϕ less than or greater than 90 degrees. From this it can be concluded that electron delocalization is responsible for the conformational behavior of $^3J_{HH'}$ of ethane.

This conclusion can be extended beyond ethane, as can be observed from the good empirical correlations reported [11] between $^3J_{HH}$ and δ(H, H′), with correlation coefficients, r^2, of 0.990 and 0.975, for several polybenzenoid and aliphatic hydrocarbons, respectively, using RHF/6-31G(*d*,*p*) wavefunctions. These results

Table 14.1 Nonlinear fit of the form $\delta(H, H') = A\cos^2\phi + B$ for δ(H, H′) for the dihedral angle, ϕ, of ethane.[a]

ϕ	A	B	Correlation coefficient
0–90	0.0067	0.0016	0.9997
90–180	0.0079	0.0018	0.9988

[a] QCISD/6-311++G(2*d*,2*p*) wavefunctions used.

support the conclusion that proton–proton vicinal coupling constants are a consequence of electron delocalization and exemplify how the Fermi exchange density contains information related to nonbonded atoms.

It is interesting to note that satisfactory empirical correlations have also been reported for F–F coupling constants [13], even though higher angular momentum functions included on the F atoms involve larger contributions from dipolar terms. This probably works because of error cancellation between other than Fermi contact terms, with the latter remaining as the most important contribution, or because of the very tight electron density in the neighborhood of the F nuclei. Further study is necessary to clarify these points.

14.3 Conformational Equilibria

14.3.1 Rotational barriers

14.3.1.1 Rotational Barrier of Ethane

Internal rotation about single bonds is discussed in many textbooks as a relevant feature of molecular conformation and its dynamics, and ethane is often considered in theoretical work as a prototype for analysis of this phenomenon. After Kemp and Pitzer [14] proposed the existence of a rotational barrier in ethane in 1937, several explanations have been invoked [15]; but the discussion of its origin is far from complete.

Of several experimental measurements reported [16, 17], a value of 2.93 kcal mol^{-1} [18] can be regarded as the most accepted value for the barrier and 2.79 kcal mol^{-1} seems to be a reasonable best estimate from CCSD(T)/6-311G(3df,2p) calculations [19]. The barrier is usually attributed to a steric effect, because of repulsion between C–H bonds or between vicinal H atoms in the eclipsed conformation [20]. A popular explanation of the barrier has been given with the use of natural bond orbital analysis in which hyperconjugation is regarded as the source of the conformational preference of the molecule [21–23] (by means of $\sigma_{C-H} \rightarrow \sigma^*_{C-H}$ vicinal interactions) although such a model has been criticized [24].

Bader et al. [25] provided an alternative explanation based on the QTAIM using the energy partitioning included in the molecular Hamiltonian without resort to any hypothetical reference state. In this partition, the contributions to the potential energy of the molecule are the attractive (nuclear–electron) and repulsive (nuclear–nuclear plus electron–electron) terms. Accordingly, the barrier results from the decrease of magnitude of the attractive interaction in the eclipsed conformation, despite the accompanying reduction in the repulsion contribution.

The origin of the rotational barrier in ethane can be traced back to the behavior of the atomic energies along the barrier, one of the main results being that the energy of a carbon atom has the same trend as the barrier (Fig. 14.2). The carbon

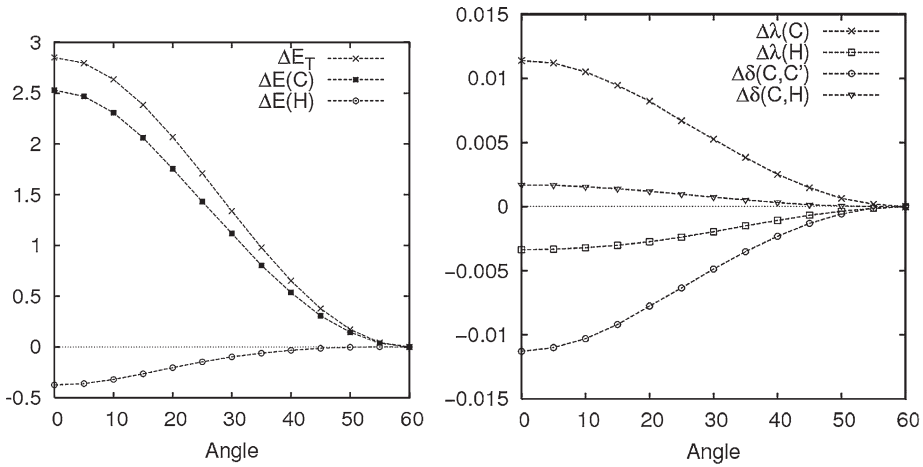


Fig. 14.2 (a) Molecular and atomic energies of carbon and hydrogen atoms along the rotational barrier of ethane. (b) Localization and delocalization indices for ethane. QCISD/6-311++G(2*d*,2*p*) wavefunctions used. Values relative to the staggered conformer.

atoms are more stable for the staggered conformer by 5.1 kcal mol^{-1} and H atoms are more stable in the eclipsed conformation with a total contribution to the barrier of −2.28 kcal mol^{-1}. The change of the potential energy of the carbon atom is mainly because of the nuclear-electron contribution of the carbon atom, +62.06 kcal mol^{-1}, which indicates a decrease in the magnitude of attraction, compared with a reduction of −57.09 kcal mol^{-1} for the atomic repulsive contribution. These values are in accordance with both the molecular results reported by Bader et al. [25] and with the Ehrenfest force – the force exerted on the electron distribution by the other electrons and the nuclei [26]. It was found there is a decrease in the attractive Ehrenfest force acting on the C atoms when rotation from the staggered to the eclipsed conformation occurs. Because in these two conformations the Feynman forces acting on the nuclei vanish, the Ehrenfest force explains the barrier. This evidence makes it unnecessary to use any repulsive model to explain the rotational barrier (for example one based on Pauli repulsions, for which no force operator can be defined) [27].

Small but decisive charge redistribution accompanies all these energy changes. The electron population of the carbon atom increases by 0.007 e (Table 14.2), as does the C–C bond distance (by 0.0141 Å) during rotation from the staggered to the eclipsed conformer. As a consequence, $\rho_b(\mathbf{r})$ at the C–C BCP decreases by 0.007 au. The electronic energy density, $H(\mathbf{r}) = G(\mathbf{r}) + V(\mathbf{r})$, the sum of the positive definite kinetic energy density plus the potential energy density, is useful for characterization of chemical bonding [28]. $H_b(\mathbf{r})$ at the C–C BCP of ethane becomes less negative by 0.011 au, indicative of lower C–C bond stability in the eclipsed conformation. Additional evidence of the charge polarization occurring

Table 14.2 Electron population, atomic energy, localization, and delocalization indices of carbon atom in ethane (au) from QCISD/6-311++G(*d*,*p*) calculations.[a]

Angle	N(C)	E(C)	λ(C)	Δ(C)	δ(C, C′)
0	5.9258	−37.87012	4.2044	1.7214	0.8422
60	5.9189	−37.87419	4.1930	1.7259	0.8535

[a] Ref. [26].

during the process is given by the shift of the charge centroid, given by minus the dipole moment of the methyl group, 0.001 au away into the nonbinding region of the molecule.

In the same process, the number of electrons located on a carbon atom increases by 0.0114 e whereas its total delocalization decreases by 0.0044 e. The main contribution to the latter, 0.0113 e, comes from the lower C–C delocalization (Fig. 14.2). In addition, the electrons a carbon shares with an hydrogen atom remain almost unchanged. The QCISD values of λ(C) are 4.2044 e and 4.1930 e for the eclipsed and staggered conformations, respectively. The main change in delocalization is that between C1 and C2, increasing in the staggered conformation by 0.0113 electron pairs (ep). The delocalization between C1 and the H atoms bonded to it decreases in the staggered conformation by 0.0017 ep whereas the C–H bonds remain almost unchanged. The delocalization between C1 and the hydrogen atoms bonded to C2 increases in the staggered conformation by 0.0009 ep. The amounts of electron localization of a carbon atom, $l = 100[\lambda(\mathrm{C1})/\mathrm{N(C1)}]$, are 70.95% and 70.84%, in each conformation. The total delocalization of carbon to the other basins, $\Delta(\mathrm{C1}) = \mathrm{N(C1)} - \lambda(\mathrm{C_1})$, is 1.7214 e for the eclipsed and 1.7259 e for the staggered conformation. The amount of electron delocalization of C1 into the other basins in the staggered conformation, $100[\delta(\mathrm{C1, X})/2\mathrm{N(C1)}]$, is 7.21% with C2 and 7.02% with each of the three hydrogen atoms bonded to C1; the rest of the delocalization, 0.88%, occurs with the hydrogen atoms bonded to C2.

It is also interesting to note that whereas Karplus-type behavior (discussed in Section 14.2) is observed for electron delocalization between vicinal hydrogen atoms, this is not so for the energy barrier. In contrast, C–C delocalization is in accordance with the features of the barrier – during rotation toward the eclipsed arrangement the electrons shared between the carbon atoms become more localized on the respective basins thus, reducing the electron–nuclear attraction and increasing the nuclear repulsion between them.

14.3.1.2 **Rotational Barrier of 1,2-Disubstituted Ethanes**

This subsection presents results from a study of the conformational behavior of the molecules $\mathrm{XCH_2CH_2X}$, where X = CH3, Cl, or F. It has been shown exper-

imentally that whereas the *antiperiplanar* (*app*) conformation is the most stable for the conformers of both butane [20] and 1,2-dichloroethane [29], the *gauche* (*sc*) conformation is the most stable for 1,2-difluoroethane [29], a behavior known as the *gauche* effect [30].

The highest of the rotational barriers of butane occurs when the two methyl groups eclipse each other and has been attributed to the action of van der Waals repulsion between them. In this instance the *anti* conformer is the one observed experimentally, with an *anti-gauche* energy difference that lies between 0.67 and 0.97 kcal mol^{-1} in the gas phase [31]. For 1,2-dichloroethane the *anti* conformer is the most stable; the energy difference between this and the *gauche* conformer is estimated to be between 0.9 and 3 kcal mol^{-1}, depending on the experimental technique used [32]. The highest of the barriers in this molecule has been attributed to dipole–dipole repulsion between the C–Cl bonds [33].

The *gauche* effect occurs when X is an electron-withdrawing substituent, for example F or O, and has been found in many molecules in addition to 1,2-disubstituted ethanes, although 1,2-difluoroethane is regarded as a typical example. Microwave, Raman, infrared, and nuclear magnetic resonance (NMR) studies indicate that the *gauche* conformer is the most stable for this molecule and the *anti–gauche* energy difference has been found to be between 0.6 and 0.9 kcal mol^{-1} [34]. The rotational barrier for 1,2-difluoroethane has been determined computationally using several levels of theory [22, 33, 35]. A value of 0.8 kcal mol^{-1} for the *anti–gauche* energy difference has been reported at the MP2/ANO level but even the RHF approximation, whenever a sufficiently large basis set is used, accounts for the *gauche* effect of the molecule [36].

Several explanations have been proposed for the *gauche* effect. On the basis of analysis of the evolution of the bent-bond C–C trajectory of the molecule during rotation, Wiberg and coworkers [37] concluded that a destabilizing interaction in the *anti* rotamer makes the *gauche* conformer the most stable. The *gauche* effect has also been explained in terms of enhanced $\sigma_{C-H} \rightarrow \sigma^*_{C-F}$ hyperconjugative interactions in the *gauche* conformer [38]. Alternatively, increased orbital overlap between the HOMO molecular orbitals of two interacting CFH_2 radicals when the F atoms are in a *gauche* arrangement has been used as an explanation [36].

As for ethane the C–C bond distance of molecules XCH_2CH_2X undergoes the most relevant changes during rotation for which, according to Table 14.3, there is a well defined trend as a function of X – it is shorter for the *anti* than for the *gauche* conformation of butane; it is nearly equal for both conformers of 1,2-dichloroethane; and is shorter for the *gauche* conformation of 1,2-difluoroethane as a consequence of the *gauche* effect.

The atomic energies of the central carbon atoms of the molecules, shown in Fig. 14.3a, indicate that whereas for butane and 1,2-dichloroethane these atoms are more stable in the *app* conformation, for 1,2-difluoroethane they are more stable at 90 degrees. In addition, the relative substituent energies shown in Fig. 14.3b indicate that X has a stabilizing effect when the substituents are eclipsed, except for Cl, for which this effect is observed when this atom becomes eclipsed with an H atom. The behavior of the substituent energy is in clear contradiction

Table 14.3 C–C bond distance in XCH_2CH_2X molecules.[a]

X	Conformer	
	syn	***gauche***
CH_3	1.530	1.535
Cl	1.510	1.508
F	1.510	1.500

[a] Values given in Å, obtained at the B3LYP/6-311++G(2*d*,2*p*) approximation.

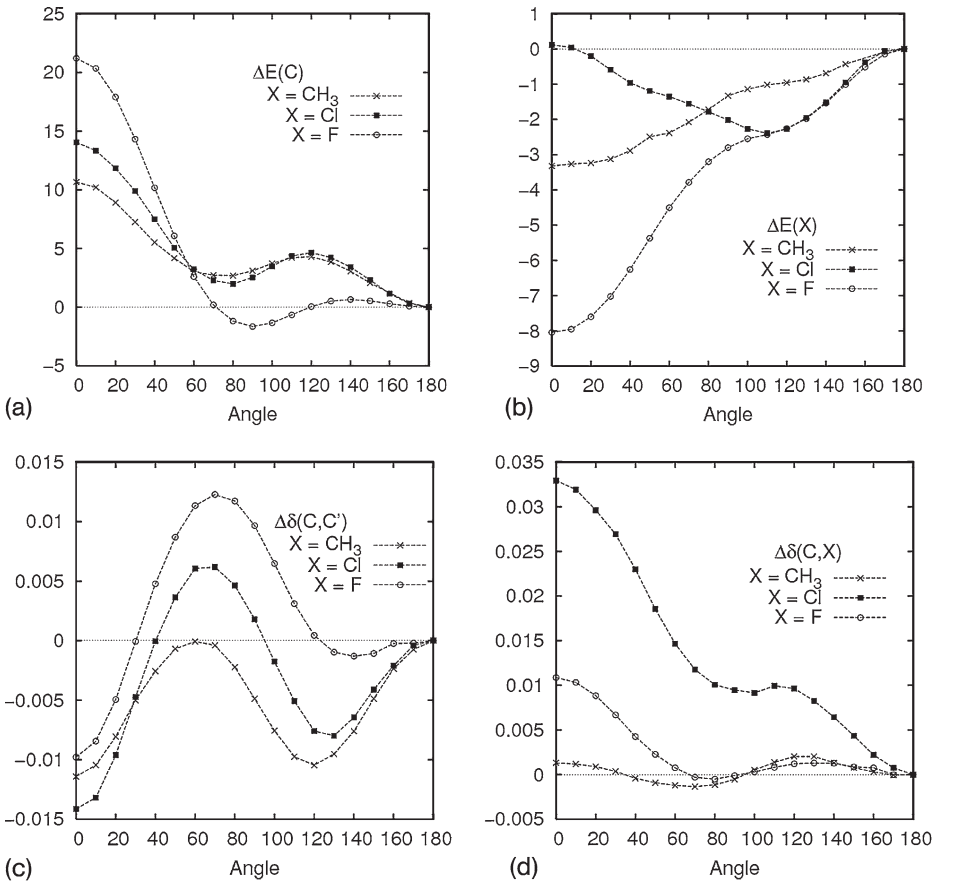


Fig. 14.3 (a) Atomic energies of the carbon atom along the rotational barrier for 1,2-disubstituted ethanes; (b) substituent energies; and (c) δ(C, C′) and (d) δ(C, X) delocalization indices. B3LYP/6-311++G (2*d*, 2*p*) wavefunctions were used. Values relative to the *anti* conformer.

with any explanation of the barrier based on atomic or functional group repulsion in the molecules; in the same way as for ethane, the main source of the barrier is the increase of E(C).

An explanation of the *gauche* effect can be provided by QTAIM analysis of the process. A lack of coincidence between the angle at which the *gauche* conformer is observed (70°) and the E(C) minimum for the three molecules is found. Whereas for butane and 1,2-dichloroethane the carbon atoms are more stable at 80°, for 1,2-difluoroethane the angle is 90°, as a result of the balance among all atomic contributions. Note that for 1,2-difluoroethane the most important energy contributions to this balance come from both the central carbon atoms and the *gauche* H atom. Of these, the former has a stabilizing and the latter a destabilizing effect; at 70°, the H atom that is not *app* to the vicinal F atom has the least destabilizing effect and the carbon atom has the most stabilizing effect. As a consequence, the *gauche* effect in 1,2-difluoroethane is observed.

Electron delocalization between the central carbon atoms follows the same trend as in ethane; δ(C, C′) is smaller for the eclipsed conformations (Figs 14.3c and 14.3d). In particular, this delocalization for the *gauche* conformer follows the order $CH_3 < Cl < F$, in agreement with the *gauche* effect for FCH_2CH_2F. C–X delocalization follows the reverse order, being larger for the eclipsed conformations. δ(F, F′) also undergoes important changes during internal rotation – it is maximum at both the F/F eclipsed and *anti* conformations and minimum for a torsion angle of ca. 100° (Fig. 14.4). δ(F, H) is smaller and maximum when F and H are at their *anti* and *syn* conformations, respectively. This does not support the classic hyperconjugative model as the origin of the *gauche* effect, however, because the change of δ(F, H) contributes only 0.006 e but the change of δ(C, C′) contributes 0.013e to the *gauche* conformer. From these results, the relative con-

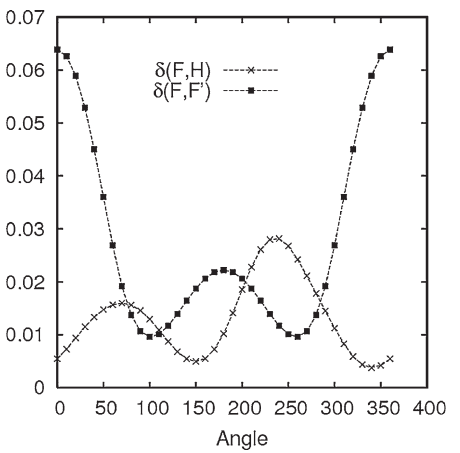


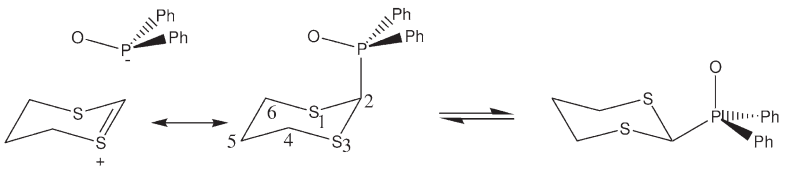
Fig. 14.4 δ(F, H) and δ(F, F′) delocalization indices. B3LYP/6-311++G (2*d*,2*p*) wavefunctions were used.

former stability can be associated with electron delocalization between the central C atoms.

14.3.2
Anomeric Effect on Heterocyclohexanes

A relevant aspect of the QTAIM to experimental chemists is the rigorous description of bonding it provides in terms of $\rho(\mathbf{r})$, because in this theory the existence of a bond path is both a necessary and sufficient condition for the existence of a chemical bond [39]; this provides chemists with an important tool for analysis of covalent, shared, and polar interactions. This subsection shows the relevance of the bond-path trajectories as a guide for the design of new compounds.

The participation of weak intramolecular interactions in a conformational process can be illustrated by the anomeric effect, defined as the thermodynamic preference of an electronegative substituent to assume the axial position when it acquires a position α to an annular heteroatom. The anomeric effect occurs in the S–C–P(O) segment when the diphenylphosphinoyl group is attached to position 2 of 1,3 dithiane (Scheme 14.1) [40]. In this example, an atypical hydrogen bridge in the CH–OP group is essential for preference of the substituent.



$\Delta G^0_{289} = 1.0$ kcal/mol
$\Delta H^0_{289} = 2.25$ kcal/mol
$\Delta S^0_{289} = 4.43$ cal/K mol

EA = 2.64 kcal/mol

Scheme 14.1 Conformational equilibrium of 2-diphenylphosphinoyl-1,3-dithiane.

In 1982 Juaristi et al. described, for the first time, the anomeric effect at the S–C–P(O) segment [41], with a value of 2.64 kcal mol^{-1}, one of the largest values yet reported for this effect. Later, the enthalpic nature of this type of effect was established [42–44]. This phenomenon baffled the scientific community – weak anomeric effects were expected because of the low electron-donating nature of the atoms involved (from the second row of the periodic table) [45, 46].

One of the models used to describe the anomeric effect relies on hyperconjugation [47]. X-ray diffraction data of the axial conformer do not, however, show the

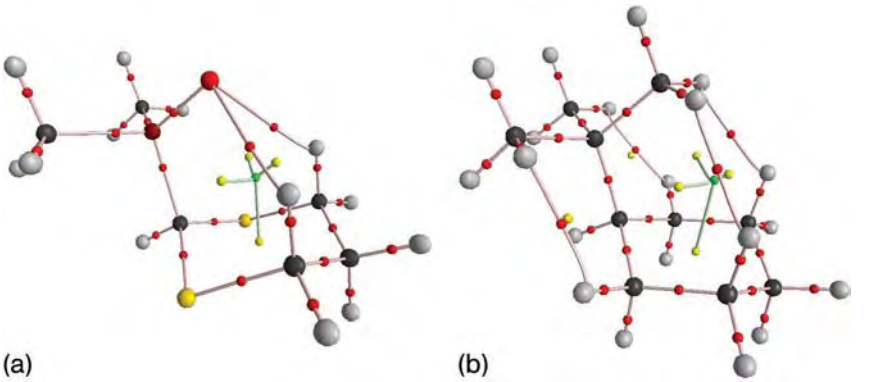


Fig. 14.5 Molecular graphs of (a) 1-ax and (b) *t*-butyl cyclohexane calculated at the B3LYP level with the 6-31G(*d*,*p*) and 6-311G(2*d*,2*p*) basis sets, respectively. Paths connecting CCPs are shown as green lines. BCPs, RCPs, and CCPs are shown as small red, yellow, and green spheres.

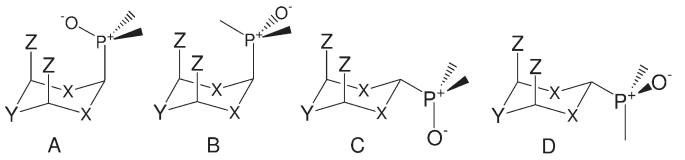
bond shortening and lengthening patterns required by this model – the C–S bond should be shorter and the C–P bond longer in the axial conformer than in the equatorial conformer, but this cannot be observed (Fig. 14.5a) [48]. An alternative explanation proposed a through-space electronic interaction between the S and P atoms. An interaction between 3p(S) and 3d(P) orbitals was then proposed. According to Graczyk [49], if the SS and SP nonbonding distances are ca 3.00 Å and the maximum radial extent of the phosphorus 3d orbital is 2.43 Å [50], either no or very small 3p–3d overlap is possible. Schleyer et al. questioned the relevance of 3p–3d interactions on theoretical grounds [51].

Mikolaczyk suggested a rationalization for the origin of the anomeric effect in the S–C–P segment in terms of an interaction between the oxygen atom of the phosphinoyl group and the hydrogen atoms at the 4,6-*syn*-diaxial positions [52]. Such a hypothesis is based on the observation that the distance between these hydrogen atoms and the oxygen atom is shorter than the sum of their van der Walls radii. On the other hand, interpretation of microwave data led Mikolajczyk et al. [53] to conclude that the hydrogen bridge on the CH–OP segment is of no relevance.

Calculation of the electronic properties of 2-dimethylphosphinoyl-1,3-dithiane (**1**) at the B3LYP/6-31G(*d*,*p*) level established that the axial conformer in which the O atom is on the ring pointing towards the axial H atoms at positions 4 and 6 is the most stable of four possible conformers (Table 14.4). Computational results obtained from experimental data establish that for evaluation of the conformational energy it is possible to replace the phenyl groups of the experimentally studied compound by methyl groups [40].

Interestingly, conformers A and B of compound **1** (see numbering scheme in Table 14.4) acquire the same atomic arrangement as that found in the S–C–P

Table 14.4 Relative energy and $H_b(\mathbf{r})$ of the CH–OP interaction of conformers A–D, in kcal mol^{-1} at B3LYP/6-31G(*d*,*p*) level.



Compound/Conformer	A	B	C	D
1: X = S, Y = CH_2, Z = H	0.0/(0.011, 0.10)	6.37	5.83	4.45
2: X = CH_2, Y = CH_2, O, Z = H	1.49/(0.009, 0.009)	3.81	0.0	0.35
3: X = S, y = S, Z = H	0.0/(0.039, 0.039)	7.61	6.38	5.01
4: X = SO_2, Y = CH_2, Z = H	0.0/(0.036, 0.036)	11.62	8.53	5.00
5: X = S, Y = CH_2, Z = F	2.96	0.39	0.0	0.21

anomeric segment, so that if any stereoelectronic effect was to participate in molecular stabilization, it should remain constant. Rotation of the phosphinoyl group results in destabilization of 6.37 kcal mol^{-1}, however. In conformer B, the methyl group points towards the center of the ring causing, the molecular energy to increase. If conformer B of compound **2** is used as a reference in which cyclohexane is taken as the basic system, an energy increase of only 3.81 kcal mol^{-1} relative to the minimum value of C is observed. This is similar to the effect experienced by a *tert*-butyl group on cyclohexane in which, as illustrated in Fig. 14.5b, the presence of hydrogen–hydrogen bond paths precludes assignment of repulsive character to 1,3-*syn* diaxial interactions in the molecule [54]. The energy increase can be a consequence of the loss of CH–OP interactions and of an increase of methyl group interactions now oriented over the dithiane ring. This would lead to a difference of 2.56 kcal mol^{-1} that can be attributed to both CH–OP interactions and stabilization energy of 1.28 kcal mol^{-1} for each of them.

Figure 14.5a shows the critical points of the molecule. The presence of two CH–OP bonds yields a molecular structure similar to that of adamantane and enables the formation of three additional rings besides that of 1,3-dithiane. A total of four ring critical points (RCPs) and one cage critical point (CCP) are therefore generated. It is interesting to note that as the two BPCs have different $\rho_b(\mathbf{r})$ values – one being stronger than the other [40]. In this instance the associated bond paths are curved; they are, therefore, longer than the geometrical bond distances. Bond-distance analysis shows the difference between the two interactions. The trajectory with $\rho_b(\mathbf{r}) = 0.011$ au corresponds to a bond length of 4.726 au whereas for the other, $\rho_b(\mathbf{r}) = 0.010$ au, corresponds to a trajectory length of 4.799 au.

Replacement of the S atoms with methylene groups, **2**, causes $\rho_b(\mathbf{r})$ to decrease [55]. Introduction of methylene groups might be expected to cause a decrease of

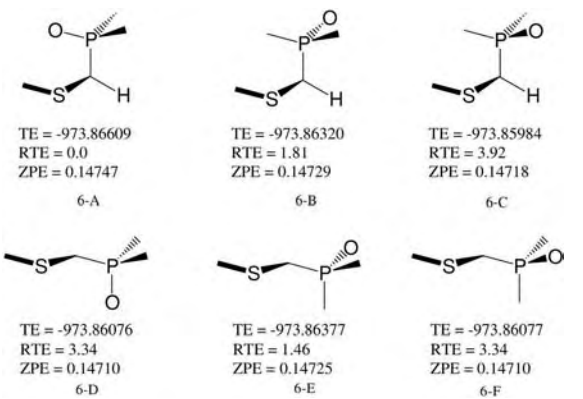
acidity of the H atoms at positions 4 and 6, a weaker CH–OP bond, and, as a consequence, the equatorial conformer should be preferred. In addition, introduction of a third S atom at position 5 of 1,3-dithiane, **3**, should increase the acidity of the C–H group, with a concomitant increase of $\rho_b(\mathbf{r})$. The experimental conformational preference of 1,3,5 trithiane is 1.43 kcal mol^{-1} [48], a larger value than that for **1**.

Something similar happens when the amount of oxidation of the sulfur atoms is increased, as shown for molecule **4**. In addition, on these molecules there is no possibility of the sulfur atoms having any stereoelectronic interaction with the P atom. The preference for the axial position is, nevertheless, preserved because the acidity of the H atoms of interest is larger than for 1,3-dithiane [55].

It is also relevant that the conformational preference of the substituent at position 2 of the axial conformation with a *gauche* arrangement of the oxygen atom increases with $\rho_b(\mathbf{r})$ for the bridge CH–OP BCP.

The strongest interaction is produced when the S atom is oxidized to its corresponding sulfone. It is, unfortunately, not possible to confirm these results experimentally for 1,3-dithiane, because of the impossibility of obtaining the disulfone while at the same time keeping the chair conformation of the 1,1,3,3-tetraoxa-1,3-dithiane ring. Because of this, dimethylphosphinoyl(methylsulfanyl)methane, **6**, was studied, taking into consideration that the acidity of the hydrogen atom that is α to S can be modulated to modify its oxidation state and, depending on the results obtained, to perform the synthesis.

Scheme 14.2 shows the six stable conformers of compound **6**. Three have a *gauche* arrangement of the C–S–C–P segment and three have the *anti* conformation. Of these, conformer **6**-A, in which the O atom approaches an H atom of the methyl group, is the most stable. The corresponding interaction energy coincides with the value obtained from analysis of 2-diphenylphosphinoyl-1,3-dithiane



Scheme 14.2 Minimum-energy conformers in the potential-energy surface of dimethylphosphinoyl(methylsulfanyl)methane (**6**). Adapted from Ref. [40].

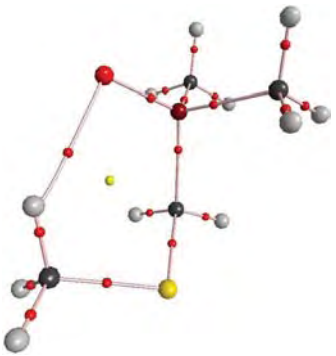


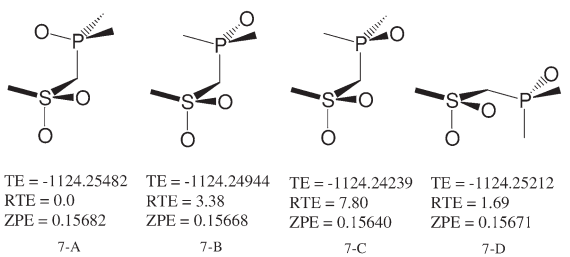
Fig. 14.6 RCPs and BCPs on conformer **6**-A. Critical are points labeled as in Fig. 14.5.

(1.8 kcal mol^{-1}). The energy of conformer **6**-C is 3.92 kcal mol^{-1} higher than that of **6**-A. The difference is less for **6**-B, because the C–S and P–O bonds are periplanar, probably because of a stabilizing $\sigma_{O-P} \rightarrow \sigma^*_{C-S}$ stereoelectronic interaction. Of the three *anti* conformers, only **6**-E (the most stable of the *gauche* series) preserves such interaction.

There is a CH–OP BCP with $\rho_b(\mathbf{r}) = 0.019$ au on conformer 6-A and an associated RCP (Fig. 14.6). The bond path length is 4.7049 au and the geometric length is 4.6524 au.

Scheme 14.3 shows the results from optimization of the geometry of dimethylphosphinoyl (methylsulfonyl)methane, 7. Only four conformers can be found, three with a *gauche* and only one with an *anti* arrangement of the C–S–C–P segment. From this, only the latter keeps an antiperiplanarity that enables $\sigma_{O-P} \rightarrow \sigma^*_{C-S}$ interactions. The energy minimum corresponds to the conformer for which the CH–OP interaction is possible. Two nonobvious BCPs are found for compound 7-A, one for the PO–HC trajectory of interest and another for the SO–CH trajectory of one of the methyl groups with a sulfonyl O atom.

The BCP typical of the CH–OP interaction can be found on **7**-A with $\rho_b(\mathbf{r}) = 0.0152$ au. A second bond path is evident for SO–HC with $\rho_b(\mathbf{r}) = 0.0111$ au.



Scheme 14.3 Conformers of minimum energy in the potential-energy surface of dimethylphosphinoyl(methylsulfonyl)methane (**7**). Adapted from Ref. [40].

That the energy difference is so large compared to the other conformers led to the proposal it be synthesized; which was accomplished with a 70% global yield from chloromethyl methyl sulfide. The NMR spectrum of **6** contain a singlet signal at 2.22 ppm for the methyl group and a doublet at 3.22 ppm, with $^2J_{H-P} = 13.0$ Hz, for the methylene group. The signals for the aromatic rings are complex and centered at ca 7.5 and 7.8 ppm. In contrast, the spectrum for compound **7** contains a triplet signal centered at 3.226 ppm and $^4J_{H-H} = 0.75$ Hz for the methyl group, and a double of quartets centered at 4.518 ppm and $^4J_{H-H} = 0.75$ y $^2J_{H-P} = 9.0$ Hz for methylene. Calculation of the coupling constants of the different hydrogen atoms of conformer 7-A leads to the conclusion a W-type coupling is responsible for the additional multiplicity. This shows that the CH–OP bridge exists and results in slower methyl rotation, thus enabling the observed $^4J_{HH}$ coupling.

It has been proposed from theoretical studies that CH–O=C interactions in purinic and pyrimidic pairs are repulsive in nature [56]. The analysis described herein enables characterization of the attractive CH–OP interaction relevant to the anomeric effect of the S–C–P–O segment, however, just as predicted by chemical intuition and emphasized theoretically [27].

14.4 Aromatic Molecules

14.4.1 Electronic Structure of Polybenzenoid Hydrocarbons

The molecular and electronic structures and intermolecular interactions of aromatic molecules have been successfully explained by the QTAIM in terms of $\rho(\mathbf{r})$, its Laplacian, $\nabla^2\rho(\mathbf{r})$, and the kinetic energy density, $H(\mathbf{r})$. Bond paths have been reported [57] between ortho H atoms in angular polybenzenoid hydrocarbons (PBHs), thus providing evidence for the presence of hydrogen–hydrogen bonding as in *tert*-butylcyclohexane [54], shown in Fig. 14.5b. Bader et al. [58] analyzed the properties of $\rho(\mathbf{r})$ of benzene and several protonated derivatives and Wiberg [59] reported good correlations among the properties of $\rho_b(\mathbf{r})$ with bond distances and with the π component of Fulton bond index of PBHs. Howard and Krygowski [60] also reported good correlations for the values of $\rho(\mathbf{r})$ at the RCPs in their study of aromaticity descriptors.

Matta et al. [61] further investigated the relationship between electronic and geometric data with the delocalization index. The correlations obtained imply that for the PBHs analyzed there is a nonobvious relationship between the one-electron information contained in $\rho_b(\mathbf{r})$ and two-electron properties such as $\delta(C, C')$ for bonded C atoms and $H_b(\mathbf{r})$. This finding allows information about the pair-density to be obtained from $\rho(\mathbf{r})$ that might not be readily available, although formally accessible, experimentally. In other words, the main features of chemical bonding found from $\delta(C, C')$ can also be obtained from the electron density of these molecules, as shown by the patterns of the electron isodensity

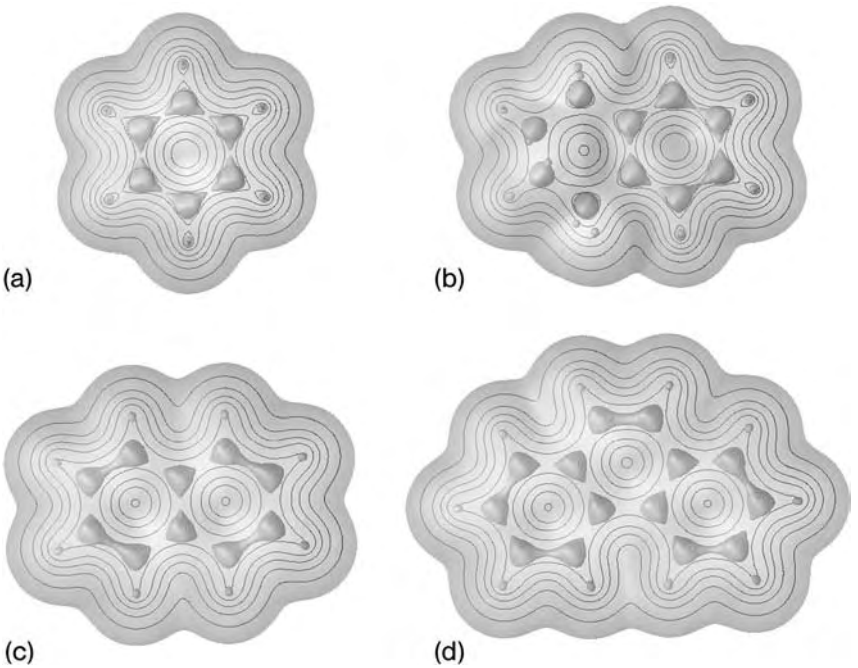


Fig. 14.7 Isodensity maps of $\rho(\mathbf{r})$ for (a) benzene, (b) tetralin, (c) naphthalene and (d) phenanthrene obtained at the HF/6-31G(*d*,*p*) approximation. Dark envelopes are for $\rho(\mathbf{r}) = 0.3274$ au, the value for the benzene C–C BCP. A value of $\rho(r) = 0.001$ au was used for the outer envelopes. Adapted from Ref. [61].

maps shown in Fig. 14.7. Notice, for example, how $\rho(\mathbf{r})$ is more localized in the inner ring of phenanthrene, as indicated by the greater volume contained in the envelope encompassing the C9–C10 bond. In addition, because more electron delocalization between bonded C atoms is associated with a more negative $\nabla^2\rho_b(\mathbf{r})$ and $H_b(\mathbf{r})$, greater bond stabilization is associated with a larger bond order quantified by $\delta(C, C')$.

All of the values reported in Table 14.5 are intermediate between those of single and double bonds. For example, the RHF/6-31G(*d*,*p*) value of $\rho_b(\mathbf{r})$ at the C–C BCP for cyclohexane is 0.255 and the corresponding $\delta(C, C')$ is 0.961, whereas for the C–C double bond of cyclohexene they are $\rho_b(\mathbf{r}) = 0.363$ and $\delta(C, C') = 1.811$. These values are to be compared with those of benzene, the reference aromatic molecule, or with the C9–C10 bond in phenanthrene, the latter of which have more double-bond character. The values reported in Table 14.5 are consistent with the envelopes of Fig. 14.7.

Aromatic monoradicals and diradicals are also of interest. For example, *m*-benzyne-derived compounds are stable toward rearrangement to their ortho and para isomers, and the experimental information suggests this molecule is in the

Table 14.5 $\rho_b(\mathbf{r})$, $\nabla^2\rho_b(\mathbf{r})$, and $H_b(\mathbf{r})$ at nonequivalent C–C BCPs and the corresponding δ(C, C′) values of several PBHs.[a]

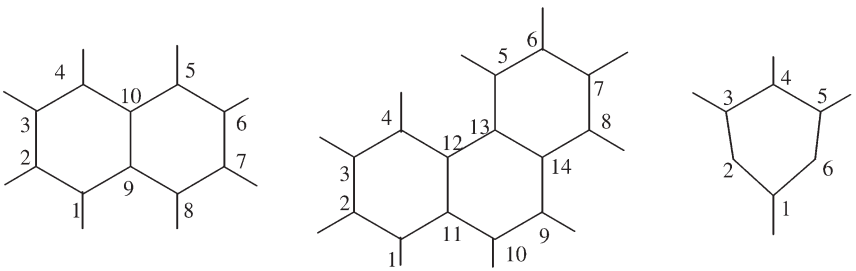
Molecule	Bond	$\rho_b(\mathbf{r})$	$\nabla^2\rho_b(\mathbf{r})$	$H_b(\mathbf{r})$	δ(C, C′)
Benzene	C–C	0.3274	−1.0163	−0.3506	1.388
Naphthalene	C1–C2	0.3432	−1.0912	−0.3855	1.533
	C1–C9	0.3095	−0.9320	−0.3128	1.222
	C2–C3	0.3106	−0.9386	−0.3161	1.245
	C9–C10	0.3170	−0.9598	−0.3273	1.278
Phenanthrene	C1–C2	0.3391	−1.0744	−0.3761	1.479
	C1–C11	0.3160	−0.9608	−0.3257	1.270
	C2–C3	0.3188	−0.9791	−0.3325	1.296
	C3–C4	0.3376	−1.0640	−0.3729	1.476
	C4–C12	0.3137	−0.9448	−0.3217	1.276
	C9–C10	0.3549	−1.1520	−0.4129	1.630
	C10–C11	0.2993	−0.8884	−0.2934	1.143
	C11–C12	0.3191	−0.9656	−0.3316	1.308
	C12–C13	0.2880	−0.8252	−0.2720	1.145
m-Benzyne	C1–C2	0.3061	−0.6579	−0.3247	1.436
	C2–C3	0.3022	−0.7956	−0.3200	1.393
	C3–C4	0.3027	−0.7955	0.2943	1.347

[a] HF/6-31G(*d*,*p*) wavefunctions used for benzene, tetralin, naphthalene, and phenanthrene [61]. In the case of *m*-benzyne in the singlet-diradical electronic state, a B3LYP/6-31++G(*d*,*p*) wavefunction within the broken symmetry approach was used [68]. The carbon atomic numbering is given in Scheme 14.4.

singlet electronic state [62], as can be inferred from its reactivity as an electrophile rather than as a free radical [63]. The importance of these compounds is illustrated, for example, by the use of *o*-benzyne nickel(0) complexes in the synthesis of PBHs [64]. Benzyne diradicals have also been investigated by theoretical methods. Nicolaides et al. [65] reported CI calculations on didehydrobenzenes and compared their results with thermochemical data [66], and Clark et al. [67] analyzed the relative stability of dehydrobenzenes to H-abstraction processes by means of CASSCF and GVB calculations. The electronic structure of aromatic radicals in the doublet state and of diradicals in both their singlet open shell and triplet states have also been investigated with QTAIM [68]. In these cases, in addition to δ(C, C′) and the properties of $\rho_b(\mathbf{r})$ at the C–C BCPs, the critical points of $\nabla^2\rho(\mathbf{r})$ have provided insight into bonding pattern for benzene and pyrene-derived radicals and diradicals – the charge concentrations revealed by the minima of $\nabla^2\rho(\mathbf{r})$ clearly indicate the radical character of phenyl and dehydropyrenes in the doublet state, and the properties of $\rho_b(\mathbf{r})$ at the C–C BCPs and the C–C′ delocalization indices are similar to the corresponding values for benzene.

More pronounced electronic redistribution occurs for singlet diradicals. For example:

1. the C–C bond properties of the dehydrogenated carbon atoms of *o*-benzyne are similar to those of a triple bond;
2. the open-shell nature of *p*-benzyne in the singlet diradical state is revealed by the existence of critical points of $\nabla^2\rho(\mathbf{r})$ on the dehydrogenated C atoms; and
3. the lack of a bond path between C1 and C5 in *m*-benzyne (Table 14.5 and Scheme 14.4) prevents one from assigning a chemical bond between these two atoms, despite the large electron delocalization between them, thus explaining the large distortion of the molecular geometry.



Scheme 14.4 Carbon-atom labeling for naphthalene, phenanthrene, and *m*-benzyne.

Similar properties have been reported for aromatic triradicals [69] by means of the QTAIM.

Although the properties of $\rho(\mathbf{r})$ at the BCP have not been regarded as useful as those at RCPs for discussion of aromaticity [60], this analysis has shown how the electronic structure of PBHs is recovered well from them and from their relationship with δ(C, C′). A useful aromaticity index can be defined [61]:

$$\theta = 1 - \frac{n}{c}\sqrt{\sum_{i=1}^{n}(\delta_0 - \delta_i)^2} \tag{12}$$

where n is the number of C atoms on a ring of a PBH, and δ_o and δ_i are the total delocalization of a carbon atom with all other carbon atoms in benzene and on the given ring of the PBH, respectively. In addition, c is a constant such that θ is zero for cyclohexane and unity for benzene and its value depends on the level of approximation used. θ does not rely on the σ–π separability (for example, the hydrogenated ring of tetralin is not planar) and can be regarded as a reformulation of the geometric HOMA [70] index taking advantage of the high correlation between C–C bond distances and δ(C, C′). θ recovers bonding characteristics of PBHs – aromatic dilution, greater electron localization and differential reactivity. This is illustrated, for example, by the following RHF/6-31G(*d*,*p*) values [61]:

$\theta = 1$ and 0.807 for benzene and naphthalene, respectively; 0.960 and 0.278 for the aromatic and hydrogenated rings of tetralin, respectively; and 0.863 and 0.650 for the outer and inner rings of phenanthrene, respectively. The bonding pattern obtained from $\rho(\mathbf{r})$, the individual C–C delocalization indices, and the θ aromaticity index, agree with the description provided by Clar's aromatic sextet model for these molecules [71]. For example:

1. the rings in naphthalene are less aromatic than in benzene; and
2. the aromaticity of the inner ring of phenanthrene is low, with the C9–C10 bond having the characteristics of a double bond, in agreement with the known addition reactivity of the molecule.

It must also be considered that further improvement of delocalization indices should account for the multidimensional nature of aromaticity with regard to the structural, electronic, magnetic and energetic aspects involved in this notion [72].

14.5 Conclusions

This chapter has shown how the QTAIM enables quantification of chemical concepts that are elusive in other approaches, for example those of the chemical bond and of aromaticity. The formality of a theory based on an observable property – the electron density – provides the experimenter with useful tools to construct models that enable explanation of their observations on key chemical concepts, for example reactivity and conformation.

It was discussed how electron delocalization is responsible for the $^3J_{HH}$ coupling constants for several systems – the Karplus-type behavior of $^3J_{HH}$ of ethane during internal rotation and trends of $^3J_{HH}$ values for cyclic aliphatic and aromatic compounds. The rotational barrier of ethane and substituted ethanes was explained on energetic grounds that are a consequence of the charge redistribution along the barrier. For example, as the Ehrenfest forces in both eclipsed and staggered conformations of ethane are attractive, but of smaller magnitude in the former arrangement, there is no need to use a repulsive model to explain the barrier. Analysis of the anomeric effect on heterocyclic molecules provides an illustrative application of how the structural criteria obtained from the QTAIM can guide the design of compounds useful for demonstrating molecular effects of interest. The electronic structure of polybenzenoid hydrocarbons obtained from the topology of the electron density and two-electron properties such as the electronic energy density and the electron delocalization index provide a clear bonding pattern and a definition of aromaticity. This broad range of examples provides a coherent description of chemical bonding and interactions in organic chemistry characterized by the QTAIM.

References

1 R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Clarendon Press, Oxford, **1995**.
2 R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic, London, **1992**.
3 E. S. Kryachko, E. V. Ludeña, *Energy Density Functional Theory of Many-Electron Systems*, Kluwer Academic Publishers, Dordrecht, **1990**.
4 R. F. W. Bader, A. Streitwieser, A. Neuhaus, K. E. Laidig, P. Speers, *J. Am. Chem. Soc.* **1996**, *118*, 4959–4965.
5 Y.-G. Wang, N. H. Werstiuk, *J. Comput. Chem.* **2003**, *24*, 379–385.
6 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem A* **1999**, *103*, 304–314.
7 T. Kar, J. A. Ángyán, A. B. Sannigrahi, *J. Phys. Chem. A* **2000**, *104*, 9953–9963.
8 R. F. W. Bader, *Coord. Chem. Rev.* **2000**, *197*, 71–94.
9 V. G. Malkin, O. L. Malkina, L. A. Eriksson, D. R. Salahub, ''The Calculation of NMR and ESR Spectroscopy Parameters using Density Functional Theory'', In: ''Modern Density Functional Theory: A Tool for Chemistry''. Eds: J. M. Seminario, P. Politzer. Theoretical and Computational Chemistry, vol. 2, Elsevier Science B. V., 1995.
10 J. A. Pople, W. G. Schneider, H. J. Bernstein, *High Resolution Nuclear Magnetic Resonance*, McGraw–Hill, New York, 1959.
11 C. F. Matta, J. Hernández-Trujillo, R. F. W. Bader, *J. Phys. Chem. A* **2002**, *106*, 7359–7375.
12 M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11–15.
13 N. Castillo, C. F. Matta, R. J. Boyd, *J. Chem. Inf. Model.* **2005**, *45*, 354–359.
14 J. D. Kemp, K. S. Pitzer, *J. Am. Chem. Soc.* **1937**, *59*, 276–279.
15 R. M. Pitzer, *Acc. Chem. Res.* **1983**, *16*, 207–210.
16 L. G. Smith, *J. Chem. Phys.* **1949**, *17*, 139–167.
17 S. Weiss, G. E. Leroi, *J. Chem. Phys.* **1968**, *48*, 962–967.
18 E. Hirota, S. Saito, Y. Endo, *J. Chem. Phys.* **1979**, *71*, 1183–1187.
19 L. Goodman, V. Pophristic, F. Weinhold, *Acc. Chem. Res.* **1999**, *32*, 983–993.
20 M. A. Murcko, H. Castejon, K. B. Wiberg, *J. Phys. Chem.* **1996**, *100*, 16162–16168.
21 L. Goodman, H. B. Gu, *J. Chem. Phys.* **1998**, *109*, 72–78.
22 L. Goodman, H. B. Gu, V. Pophristic, *J. Chem. Phys.* **1999**, *110*, 4268–4275.
23 F. Weinhold, *Angew. Chem. Int. Ed.* **2003**, *42*, 4188–4194.
24 F. M. Bickelhaupt, E. J. Baerends, *Angew. Chem. Int. Ed.* **2003**, *42*, 4183–4188.
25 R. F. W. Bader, J. R. Cheesman, K. E. Laidig, K. B. Wiberg, C. Breneman, *J. Am. Chem. Soc.* **1990**, *112*, 6530–6536.
26 F. Cortés-Guzmán, J. Hernández-Trujillo, G. Cuevas, unpublished results.
27 R. F. W. Bader, *Chem. Eur. J.* **2006**, *12*, 2896–2905.
28 D. Cremer, E. Kraka, *Angew. Chem. Int. Ed.* **1984**, *23*, 627–628.
29 K. B. Wiberg, M. A. Murcko, *J. Phys. Chem.* **1987**, *91*, 3616–3620.
30 N. C. Craig, A. Chen, K. H. Suh, S. Klee, G. C. Mellau, B. P. Winnewisser, M. Winnewisser, *J. Am. Chem. Soc.* **1997**, *119*, 4789–4790.
31 W. A. Herrebout, B. J. van der Veken, A. Wang, J. R. Durig, *J. Phys. Chem.* **1995**, *99*, 578–585.
32 P. Felder, H. H. Gunthard, *Spectrochim. Acta* **1980**, *36A*, 223–224.
33 E. L. Eliel, S. H. Wilen, *Stereochemistry of Organic Compounds*, Wiley: **1994**, 597–664.
34 D. A. Dixon, B. E. Smart, *J. Phys. Chem.* **1988**, *92*, 2729–2733.
35 K. B. Wiberg, M. A. Murcko, K. E. Laidig, P. J. McDougall, *J. Phys Chem.* **1990**, *94*, 6956–6959.

36 O. Engkvist, G. Karlström, P.-O. Widmark, *Chem. Phys. Lett.* **1997**, *265*, 19–23.

37 K. B. Wiberg, *Acc. Chem. Res.* **1996**, *29*, 229–234.

38 A. J. Kirby, *Stereoelectronic Effects*, Oxford Science Publications, **1996**.

39 R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102*, 7314–7323.

40 G. Cuevas, *J. Am. Chem. Soc.* **2000**, *122*, 692–698.

41 E. Juaristi, L. Valle, C. Mora-Uzeta, B. A. Valenzuela, P. Joseph-Nathan, M. F. Fredrich, *J. Org. Chem.* **1982**, *47*, 5038.

42 E. Juaristi, G. Cuevas, *Tetrahedron Lett.* **1992**, *33*, 2271–2274.

43 E. Juaristi, G. Cuevas, *J. Am. Chem. Soc.* **1993**, *115*, 1313–1316.

44 F. Cortés, J. Tenorio, O. Collera, G. Cuevas, *J. Org. Chem.* **2001**, *66*, 2918–2924.

45 P. v. R. Schleyer, E. D. Jemmis, G. W. Spitznagel, *J. Am. Chem. Soc.* **1985**, *107*, 6393–6394.

46 E. Juaristi, G. Cuevas, ''Recent Studies of the Anomeric Effect'', Tetrahedron Report No. 315, *Tetrahedron* **1992**, *48*, 5019–5087.

47 E. Juaristi, G. Cuevas, *The Anomeric Effect*, CRC Press. Boca Ratón Fl. **1994**.

48 E. Juaristi, L. Valle, B. A. Valenzuela, M. A. S. C. P. Aguilar, *J. Am. Chem. Soc.* **1986**, *108*, 2000–2005.

49 P. P. Graczyk, M. Mikolajczyk, *Topics Stereochem. Anomeric Effect: Origin and Consequences*, **1994**, 24, p. 277.

50 C. Yang, E. Goldstein, S. Breffle, S. Jin, *J. Mol. Struct. (Theochem)*, **1992**, *259*, 345–368.

51 A. E. Reed, P. v. R. Schleyer, *Inorg. Chem.* **1988**, *27*, 3969–3987.

52 M. Mikolajczy, *Pure Appl. Chem.* **1987**, *59*, 983.

53 P. P. Graczyk, M. Mikolajczyk, M. Pals, A. Kolbe, *J. Mol. Struct.* **1997**, *416*, 179–185.

54 F. Cortés-Guzmán, J. Hernández-Trujillo, G. Cuevas, *J. Phys. Chem. A* **2003**, *107*, 9253–9256.

55 G. Madrid, A. Rochín, E. Juaristi, G. Cuevas, *J. Org. Chem.* **2001**, *66*, 2925–2931.

56 C. Fonseca Celia, F. M. Bickelhaupt, J.-G. Snijders, E. J. Baerends, *Chem. Eur. J.* **1999**, *5*, 3581–3594.

57 C. F. Matta, J. Hernández-Trujillo, R. F. W. Bader, *Chem. Eur. J.* **2003**, *9*, 1940–1951.

58 R. F. W. Bader, C. Chang, *J. Phys. Chem.* **1989**, *93*, 5095–5107.

59 K. B. Wiberg, *J. Org. Chem.* **1997**, *62*, 5720–5727.

60 S. T. Howard, T. M. Krygowski, *Can. J. Chem.* **1997**, *75*, 1174–1181.

61 C. F. Matta, J. Hernández-Trujillo, *J. Phys. Chem. A* **2003**, *107*, 7496–7504; correction: C. F. Matta, J. Hernandez-Trujillo, *J. Phys. Chem. A* **2005**, *109*, 10798.

62 K. K. Thoen, H. I. Kenttämaa, *J. Am. Chem. Soc.* **1997**, *119*, 3832–3833.

63 E. D. Nelson, A. Artau, J. M. Price, H. I. Kenttämaa, *J. Am. Chem. Soc.* **2000**, *122*, 8781–8782.

64 K. R. Deaton, M. S. Gin, *Org. Lett.* **2003**, *5*, 2477–2480.

65 A. Nicolaides, W. T. Borden, *J. Am. Chem. Soc.* **1993**, *115*, 11951–11957.

66 P. G. Wenthold, R. R. Squires, *J. Am. Chem. Soc.* **1994**, *116*, 6401–6412.

67 A. E. Clark, E. R. Davidson, *J. Am. Chem. Soc.* **2001**, *123*, 10691–10698.

68 J. Hernández-Trujillo, I. García-Cruz, J. M. Martínez-Magadán, *Chem. Phys.* **2005**, *308*, 181–192.

69 H. M. T. Nguyen, T. Höltz, G. Gopakumar, T. Veszprémi, J. Peeters, M. T. Nguyen, *Chem. Phys.* **2005**, *316*, 125–140.

70 J. Kruszewski, T. M. Krygowski, *Tetrahedron Lett.* **1972**, 3839–3842.

71 E. Clar, *The Aromatic Sextex*, John Wiley and Sons Ltd., London, **1972**.

72 V. I. Minkin, M. N. Glukhovtsev, B. Y. Simkin, *Aromaticity and Antiaromaticity: Electronic and Structural Aspects*, Wiley, New York, **1994**.

15
Aromaticity Analysis by Means of the Quantum Theory of Atoms in Molecules

Eduard Matito, Jordi Poater, and Miquel Solà

15.1
Introduction

Aromaticity is a concept formulated to account for the unusual properties of an important class of organic and inorganic molecules – the aromatic compounds [1, 2]. Aromaticity is currently enjoying a resurgence of interest as its scope of application has expanded from the organic to the inorganic realm of chemistry [3]. The new class of all-metal and inorganic aromatic compounds which has been synthesized has prompted new definitions of aromaticity applicable to both classic and novel aromatic molecules. According to the most recent definition of Schleyer and coworkers [4], aromaticity is a manifestation of electron delocalization in closed circuits, in either two or three dimensions, which results in energy reduction, often quite substantial, and a variety of unusual chemical and physical properties. These include a tendency toward bond length equalization, unusual reactivity, characteristic spectroscopic features, and distinctive magnetic properties related to the strong induced ring currents in aromatic systems [5].

Although aromaticity is not a directly observable quantity, its importance as a central concept in chemistry has not diminished since the discovery of benzene by Michael Faraday in 1825 [6]. Yet its quantification has proved remarkably elusive. Nearly everyone agrees there is not yet a well-established method for quantifying the aromatic character of molecules. Aromaticity is usually evaluated indirectly by measuring a physicochemical property that reflects the aromatic character of molecules. Thus, most aromaticity indicators are based on the classical aromaticity criteria, namely, structural, magnetic, energetic, and reactivity-based measures [2], although, more recently, new ways of quantifying aromaticity based on the electronic properties of the molecules have been devised (a review has recently been published [7]). Although the existence of many aromaticity descriptors complicates matters, it is also true that the use of differently-based aromaticity criteria is recommended for aromaticity analysis, because of its multidimensional character [8, 9].

Structure-based measures of aromaticity rely on the idea that important manifestations of aromaticity are equalization of bond lengths and symmetry. Among the most common structure-based indices of aromaticity, one of the most effective is the harmonic oscillator model of aromaticity (HOMA) index [10], defined by Kruszewski and Krygowski as:

$$HOMA = 1 - \frac{\alpha}{n}\sum_{i=1}^{n}(R_{opt} - R_i)^2, \tag{1}$$

where n is the number of bonds considered and α is an empirical constant fixed to give HOMA = 0 for a model nonaromatic system and HOMA = 1 for a system with all bonds equal to an optimum value R_{opt}, assumed to be achieved for fully aromatic systems. R_i stands for a running bond length. The HOMA value can be split into the energetic (EN) and geometric (GEO) contributions according to the relationship [11]:

$$HOMA = 1 - EN - GEO = 1 - \alpha(R_{opt} - \bar{R})^2 - \frac{\alpha}{n}\sum_{i}(\bar{R} - R_i)^2. \tag{2}$$

The GEO contribution measures the decrease/increase in bond length alternation and the EN term takes into account the lengthening/shortening of the mean bond lengths ($\bar{R}$) of the ring. The higher the HOMA value the more aromatic the system.

Magnetic indices of aromaticity are based on the π-electron ring current that is induced when the system is exposed to external magnetic fields. Probably the most widely used magnetic-based indicator of aromaticity is the nucleus-independent chemical shift (NICS) proposed by Schleyer and coworkers [4, 12]. It is defined as the negative value of the absolute shielding computed at a ring center or at some other interesting point of the system. Rings with large negative NICS values are regarded as aromatic. The more negative the NICS values, the more aromatic the rings. Antiaromatic rings, in contrast, are characterized by positive values of NICS.

Finally, energy-based indices of aromaticity make use of the fact that conjugated cyclic π-electron compounds are more stable than their chain analogues. The most commonly used energetic measure of aromaticity is the aromatic stabilization energy (ASE), calculated as the reaction energy of a homodesmotic reaction [13].

Although aromaticity plays a prominent role in current chemistry, aromaticity measures based on electronic descriptors are rare. Among these, we can mention the HOMO–LUMO gap, absolute and relative hardness, the electrostatic potential, and the polarizability [14]. None refers directly to the electronic delocalization in aromatic species. If we take into account that electron delocalization is the

main factor responsible for manifestation of aromaticity, this fact is surprising. The electron structure of the molecules is obviously strongly connected with aromaticity, and QTAIM provides a wide set of concepts to characterize electronic structure – atomic populations and charges, electron localization and delocalization (through the so-called LIs and DIs, vide infra), BCP, RCP (one should bear in mind that aromaticity manifests in molecules with ring structures), among others [15]. In this sense, the QTAIM provides a formidable scheme for the construction of measures of aromaticity.

In the last five years we have concentrated our research efforts on quantifying aromaticity by measuring the extent of electron delocalization in molecules using the set of tools that the QTAIM theory provides for this task [7]. In this chapter we briefly review the new aromaticity criteria defined by us that make use of the QTAIM concepts. The chapter is organized as follows. Section 15.2 covers the definition of electron delocalization in the QTAIM theory; Section 15.3 analyzes electron delocalization in benzene, the quintessential aromatic molecule; Section 15.4 gives the definition of the new aromatic indexes based on analysis of electron delocalization; Section 15.5 discusses several applications of the newly defined indexes to the analysis of aromaticity; and, finally, in Section 15.6, we summarize the main conclusions.

15.2
The Fermi Hole and the Delocalization Index

Most methods used to determine the extent of electronic localization/delocalization in molecules employ the two-electron density, $\Gamma(\vec{r}_1\sigma_1, \vec{r}_2\sigma_2)$, also named second-order density or pair density [16]. This is the simplest quantity that describes the pair behavior and is usually interpreted as the probability density of finding two electrons with spins σ_1 and σ_2 simultaneously at positions $\vec{r}_1$ and $\vec{r}_2$, respectively, irrespective the position and spin of the other $N-2$ electrons. Integration of this function over the spin variables yields the spin-less pair density, $\Gamma(\vec{r}_1, \vec{r}_2)$, which can be split into an uncorrelated pair density and a part that gathers all exchange and correlation effects:

$$\sum_{\sigma_1,\sigma_2} \Gamma(\vec{r}_1\sigma_1, \vec{r}_2\sigma_2) = \Gamma(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1)\rho(\vec{r}_2) + \Gamma_{xc}(\vec{r}_1, \vec{r}_2). \tag{3}$$

The uncorrelated component of the pair density, given by the product $\rho(\vec{r}_1)\rho(\vec{r}_2)$, provides the probability of finding simultaneously two *independent* electrons at positions $\vec{r}_1$ and $\vec{r}_2$. The difference between $\Gamma(\vec{r}_1, \vec{r}_2)$ and $\rho(\vec{r}_1)\rho(\vec{r}_2)$ is known as the exchange-correlation density [17], $\Gamma_{xc}(\vec{r}_1, \vec{r}_2)$, which is a measure of the extent to which density is excluded at $\vec{r}_2$ because of the presence of an electron at $\vec{r}_1$. Integration of the exchange-correlation density of a given molecule

through all space yields the negative of the total number of electrons in this molecule.

The Fermi and Coulomb hole density is another two-electron function directly connected to the exchange-correlation density. It is defined as:

$$\rho_{xc}(\vec{r}_1, \vec{r}_2) = \frac{\Gamma_{xc}(\vec{r}_1, \vec{r}_2)}{\rho(\vec{r}_1)} \tag{4}$$

and represents the decrease/increase, relative to the uncorrelated probability density, in the probability of finding an electron in position $\vec{r}_2$ when a reference electron is fixed at position $\vec{r}_1$. In two landmark works [18, 19], Bader and Stephens using a Hartree–Fock (HF) wavefunction, which includes only the Fermi hole, showed that the extent of localization or delocalization of an electron at $\vec{r}_1$ is determined by the corresponding localization/delocalization of its Fermi hole. A localized Fermi hole indicates the presence of localized electronic charge in the position of the electron of reference and vice versa. For this reason, Fermi density hole maps have been widely used to analyze the electronic localization/delocalization [20].

Equation (4) shows that the exchange-correlation density is nothing but the Fermi and Coulomb hole density weighted by the density of the reference electron. Bader and coworkers [19, 21] used this quantity to define the localization and delocalization indexes (LIs and DIs) from the double integration of the exchange-correlation density over the atomic basins defined within the QTAIM theory:

$$\lambda(A) = -\int_A \int_A \Gamma_{xc}(\vec{r}_1, \vec{r}_2)\, d\vec{r}_1\, d\vec{r}_2 \tag{5}$$

$$\delta(A, B) = -2\int_B \int_A \Gamma_{xc}(\vec{r}_1, \vec{r}_2)\, d\vec{r}_1\, d\vec{r}_2 \tag{6}$$

The term $\delta(A, B)$ is a quantitative measure of the number of electrons delocalized or shared between atomic basins A and B, and $\lambda(A)$ is a measure of the average number of electrons localized on basin A. The following sum rule can easily be demonstrated:

$$N(A) = \lambda(A) + \frac{1}{2}\sum_{B \neq A} \delta(A, B) \tag{7}$$

Equation (7) proves that the total number of electrons belonging to a given basin can be exactly partitioned into its localized $(\lambda(A))$ and delocalized $\left(\frac{1}{2}\sum_{B \neq A} \delta(A, B)\right)$ parts. In addition, for closed-shell wavefunctions, one can define the global delocalization or valence of atomic basin A as:

$$V(A) = \sum_{B \neq A} \delta(A, B) = 2[N(A) - \lambda(A)] \quad (8)$$

As shown in the next sections of this chapter, we have used these quantities to define new electronically based aromaticity descriptors.

15.3 Electron Delocalization in Aromatic Systems

In 1996, Bader and coworkers [22], in a seminal work, analyzed electron delocalization in benzene, the archetypal aromatic molecule. They observed by means of contour maps of the Fermi-hole density in conjugated species that the interatomic delocalization of the π-electrons on a given atom, in general, decreases with the distance of the second atom from the one in question. For benzene, however, there is significantly more delocalization of the π density into the basins of the *para*-related carbon atoms than into those of the *meta*-related carbon atoms, despite the shorter distance to the latter. These data are in accordance with the energy ordering of the principal resonance structures of benzene – the two Kekulé structures are the most important, then the Dewar structure connecting *para*-related carbon atoms, which is, in turn, more relevant than that connecting *meta*-related atoms. To corroborate what happens in benzene, Bader and coworkers [22] studied the effect of geometrical distortion on delocalization of the π electrons in benzene by considering a symmetrical distortion (S) in which each equilibrium C–C bond length of 1.42 Å was increased by 0.06 Å in an a_{1g} stretching mode, and a b_{2u} unsymmetrical one (U) obtained by alternately increasing and reducing the bond lengths to 1.54 and 1.34 Å, respectively. The results indicated there is no significant change in the delocalization of the π electrons for S; for U, however, it is seen how delocalization between *para* carbon atoms largely decreases. The values of the DIs for benzene also confirm the larger delocalization in the *para* form than in the *meta* form. In particular, Bader and coworkers obtained the following DIs at the HF/6-31G(d) level of theory: $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{para}} = 0.101$ e and $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{meta}} = 0.070$ e. At the same level of theory, it is important to recognize that Fulton and Mixon reported, some years previously, almost identical values for $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{para}}$ and $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{meta}}$ [23]. Interestingly, these authors showed that the $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{para}}$ has a large π component (0.09 e), at variance with $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{meta}}$. The larger DI found between *para*-related carbon atoms than between *meta*-related carbons has been corroborated by use of larger basis sets and higher levels of calculation. For example, the CISD/6-311G(d,p) values of $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{para}}$ and $\delta(\mathrm{C}, \mathrm{C}')_{\mathrm{meta}}$ in benzene are 0.071 and 0.054 e, respectively [24]. This result is, therefore, not an artifact of the method used and, consequently, it has sound physical foundation for being the basis of the definition of our *para*-DI (PDI) index of aromaticity (vide infra).

15.4
Aromaticity Electronic Criteria Based on QTAIM

Sondheimer defined as aromatic those molecules with a "*measurable degree of cyclic delocalization of a π-electron system*" [25]. Likewise, Schleyer and coworkers [26] considered the aromaticity "*associated with cyclic arrays of mobile electrons with favorable symmetries*" and "the *unfavorable symmetry properties of antiaromatic systems lead to localized, rather than to delocalized electronic structures*". These definitions stress the existence of a direct connection between aromaticity and electron delocalization. As said in the introduction, unfortunately, aromaticity does not have a universally accepted quantitative descriptor. This is the reason for the continuous search for new aromaticity indexes, and the exhaustive revision of the existing ones in this quest for a less ambiguous index which, at the same time, agrees with the most elementary chemical basis of aromaticity. Two aromaticity measures based on DIs have recently been proposed, the PDI and the fluctuation aromatic index (FLU) measures. In this context it is worth mentioning, first, the work of Matta and Hernández-Trujillo [27] who attempted to construct an HOMA-like index from the QTAIM by substituting the bond length by the *total* electron delocalization, and, second, the use by Bultinck et al. [28] of the *n*-center electron DIs as descriptors of aromaticity.

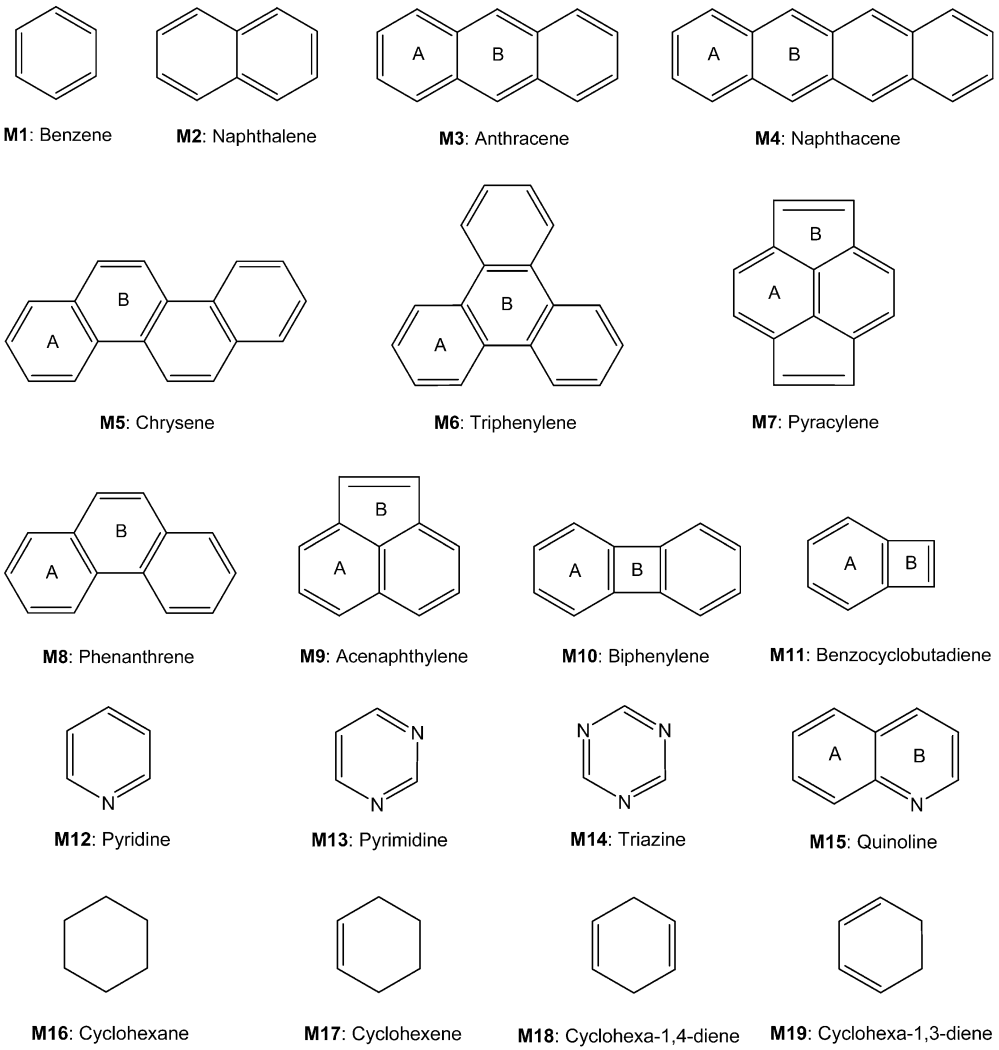
15.4.1
The *para*-Delocalization Index (PDI)

As already said, Bader and coworkers reported that delocalization in benzene is greater for *para*-related (*para*-DI) than for *meta*-related carbon atoms (*meta*-DI). With this idea in mind, Poater et al. [29] decided to undertake a study to validate the averaged *para*-DI in six-membered rings (6-MRs) as a measure of local aromaticity. The PDI is a specific measure of aromaticity for 6-MR, in which there are three *para*-related positions, namely (1, 4), (2, 5) and (3, 6):

$$PDI = \frac{\delta(1,4) + \delta(2,5) + \delta(3,6)}{3} \tag{9}$$

where $\delta(A, B)$, the DI, is defined as in Eq. (6).

The correlation between PDI and other measures of aromaticity (HOMA, magnetic susceptibility, and NICS) for the series of **M1** to **M10** polycyclic aromatic hydrocarbons (PAHs) given in Scheme 15.1 served to validate this measure as a reliable index of aromaticity [29]. The above cited work by Matta and Hernández-Trujillo [27] appeared shortly after the introduction of the PDI with the proposal of another local aromaticity index, θ, similar to the geometric HOMA, but using the DI as a measure of electron sharing alternation within a ring. This index θ was calculated for a series of PAHs revealing perfect correlation with HOMA, and a relatively good correlation with NICS. It has been also found that PDI and the index θ result in the same trends for a series of common aromatic molecules [7].



Scheme 15.1 Reprinted, with permission, from Ref. [31]; copyright 2005, American Institute of Physics.

The main shortcoming of PDI is that it can only be applied to 6-MR. This disadvantage was somehow overcome by the introduction of ΔDI [29]. ΔDI is based on the intuitive idea of comparing the electron delocalization between formal single and double bonds. For compounds with a formal Lewis structure, the difference between the DI for double and single bonds in a given ring was suggested as a measure of aromaticity. The smaller the difference (the lower the value of ΔDI) the closer is electron delocalization in double and single bounds, an indication that both single and double bonds have delocalization typical of aromatic com-

pounds. The greater the difference, on the other hand, the closer the structure to the Lewis structure, indicative of quite localized electrons, which is known to prevent aromaticity. This relationship was especially useful for five-membered ring (5-MR) species, for example the series C_4H_4–X (X = CH^-, NH, S, O, SiH^-, PH, CH_2, AlH, SiH^+, BH, CH^+), for which a reasonable agreement between ΔDI and NICS values was obtained. Nonetheless, ΔDI needs a clear Lewis structure to compare double and single bonds, and its application to rings of different sizes is less clear. In addition, nonaromatic and antiaromatic systems are not well differentiated by ΔDI values. In this sense, the quest for a new aromaticity index based on DIs, powerful enough to deal with rings of any size and able to distinguish between nonaromatic and antiaromatic species, led to the aromatic fluctuation index (FLU).

15.4.2
The Aromatic Fluctuation Index (FLU)

Sondheimer's definition [25] suggests using cyclic electron delocalization as a measure of aromaticity. Although PDI focused on *para*-related carbon atoms and ΔDI on a couple of bonds in the ring, there was no attempt to construct an aromaticity index by examining the DI of all bonded pairs in a given ring. It is worth mentioning, however, that Bird [30] has compared Gordy bond orders of all bonded pairs in a given ring to define a measure of aromaticity that has some resemblance to our FLU index. These two indexes are similar in the sense that both compare the values of the bond orders for all pairs of adjacent atoms in the ring with a value of reference to give a measure of aromaticity. The FLU index was constructed by following the HOMA philosophy, i.e. measuring divergences (DI differences for each single pair bonded) from aromatic molecules chosen as a reference; cf. Eq. (1). The formula below was given for FLU [31]:

$$FLU = \frac{1}{n}\sum_{A-B}^{RING}\left[\left(\frac{V(B)}{V(A)}\right)^{\alpha}\left(\frac{\delta(A,B)-\delta_{ref}(A,B)}{\delta_{ref}(A,B)}\right)\right]^2 \tag{10}$$

where the summation runs over all adjacent pairs of atoms around the ring, n is equal to the number of atoms of the ring, $V(A)$ is the global delocalization of atom A given in Eq. (8), $\delta(A,B)$ and $\delta_{ref}(A,B)$ are the DI values for the atomic pairs A and B and its reference value, respectively, and

$$\alpha = \begin{cases} 1 & V(B) > V(A) \\ -1 & V(B) \le V(A) \end{cases} \tag{11}$$

The second factor in Eq. (10) measures the relative divergence with respect to a typical aromatic system, and the first factor in Eq. (10) penalizes those with highly localized electrons. The reference DI values for C–C and C–N bonds were obtained from benzene (1.4 e) and pyridine (1.2 e) at the HF/6-31G(d) level of

theory. In a forthcoming work we will also give the reference data for B–N bonds, which will be taken from (0.77 e) [32].

As is readily apparent, unlike PDI, FLU can cope with rings of different sizes and furnishes global and local measures of aromaticity. The weakness of FLU, however, as it is also for HOMA, is the need for typical aromatic systems as references. One must also be aware that FLU is actually measuring the relative electronic divergence of a given ring in a molecule relative to a molecule chosen as a reference. It is, therefore, worth noting that FLU (the same is true for HOMA) must be applied with care when studying the change of aromaticity along a reaction coordinate, because it fails to recognize instances when aromaticity is enhanced on deviation from the equilibrium geometry (vide infra) [33].

15.4.3
The π-Fluctuation Aromatic Index (FLU_π)

To overcome the need for reference data which prevents the FLU to be applied in a straightforward manner to any molecule, another index based on the QTAIM was designed – the FLU_π; this measures the divergence of π-delocalization from its average [31]:

$$FLU_\pi = \frac{1}{n}\sum_{A-B}^{RING}\left[\left(\frac{V_\pi(B)}{V_\pi(A)}\right)^\alpha\left(\frac{\delta_\pi(A,B)-\delta_{av}}{\delta_{av}}\right)\right]^2 \tag{12}$$

where δ_{av} is the average value of the π-DI for the bonded pairs in the ring, and the other symbols denote the aforementioned quantities calculated using π-orbitals only. FLU_π can only be exactly calculated for planar molecules, where DI can be (exactly) decomposed into its σ and π contributions. Nevertheless, orbital localization schemes can always be used to obtain approximate π-DI for nonplanar molecules.

The correlation between FLU_π and FLU is shown to be excellent for a series of organic compounds (cf. Table 15.1 and Scheme 15.1). There is, however, no reason to expect this always to be true. Indeed, FLU_π is measuring the amount of homogeneous delocalization in a π-system, whereas FLU is measuring the extent of similarity with reference aromatic molecules. Hence, for organic species for which aromaticity comes from delocalization of the π-system, one expects both indexes to reproduce the same trends; differences will, however, arise for inorganic species, the aromaticity of which is not exclusively driven by π-electron delocalization.

Correlation of PDI, FLU, and FLU_π with other aromaticity indexes is usually reasonably good (cf. Table 15.2) for the series of compounds given in Scheme 15.1, perhaps with the exception of NICS values, as one would expect. One can, therefore, see how QTAIM provides a formidable scheme for measurement of aromaticity. It is also worth noticing that some other molecular decompositions, for example fuzzy-atom [34], can furnish measures of aromaticity in excellent agreement with those derived from QTAIM DIs [35].

	Ring	EN	GEO	HOMA	NICS(0)	$NICS(0)_{zz}$	NICS(1)	$NICS(1)_{zz}$	PDI	FLU	FLU_{π}
Benzene		−0.001	0.000	1.001	−11.5	−17.2	−12.9	−32.5	0.105	0.000	0.000
Naphthalene		0.022	0.199	0.779	−10.9	−13.8	−12.6	−30.6	0.073	0.012	0.116
Anthracene	A	0.067	0.416	0.517	−8.7	−6.9	−10.7	−24.8	0.059	0.024	0.254
	B	0.045	0.072	0.884	−14.2	−21.9	−15.5	−38.4	0.070	0.007	0.024
Naphthacene	A	0.110	0.565	0.325	−6.7	−0.9	−9.0	−19.6	0.051	0.031	0.355
	B	0.073	0.153	0.774	−13.8	−20.4	−15.2	−37.2	0.063	0.011	0.073
Chrysene	A	0.011	0.130	0.859	−11.1	−13.7	−12.8	−30.7	0.079	0.008	0.068
	B	0.122	0.325	0.553	−8.2	−3.6	−10.6	−23.3	0.052	0.019	0.185
Triphenylene	A	0.003	0.068	0.930	−10.6	−11.8	−12.3	−29.3	0.086	0.003	0.026
	B	0.609	0.324	0.067	−2.6	13.9	−6.2	−9.4	0.025	0.027	0.181
Pyracylene	A	0.000	0.329	0.671	−4.9	6.9	−7.2	−13.3	0.067	0.014	0.132
	B	0.406	0.923	−0.328	10.1	56.4	4.8	24.2	[b]	0.050	0.686
Phenanthrene	A	0.007	0.091	0.902	−11.4	−14.7	−13.0	−31.5	0.082	0.005	0.045
	B	0.189	0.409	0.402	−6.8	−0.1	−9.4	−20.0	0.053	0.025	0.257
Acenaphthylene	A	0.011	0.192	0.797	−9.6	−8.7	−11.4	−26.4	0.070	0.013	0.117
	B	0.337	0.703	−0.039	2.2	33.7	−1.8	4.6	[b]	0.045	0.587
Biphenylene	A	0.000	0.193	0.807	−6.7	−0.7	−8.1	−17.3	0.088	0.008	0.068
	B	1.360	0.571	−0.930	17.4	87.0	7.5	35.7	[b]	0.048	0.297
Benzocyclobutadiene	A	0.001	0.501	0.497	−4.0	8.4	−5.4	−9.5	0.080	0.022	0.196
	B	0.910	1.526	−1.437	20.2	96.2	10.7	44.1	[b]	0.071	1.072
Pyridine		−0.006	0.001	1.005	−9.5	−15.2	−12.5	−31.6	0.097	0.001	0.001
Pyrimidine		0.015	0.000	0.985	−7.5	−11.8	−11.7	−29.8	0.089	0.005	0.003
Triazine		0.023	0.000	0.977	−5.3	−6.9	−10.8	−27.0	0.075	0.013	0.000
Quinoline	A	0.018	0.190	0.792	−11.0	−14.7	−12.6	−30.7	0.072	0.015	0.126
	B	0.008	0.161	0.830	−9.1	−11.5	−12.1	−29.5	0.071	0.017	0.129
Cyclohexane		5.340	0.000	−4.340	−2.1	23.5	−2.0	3.1	0.007	0.091	[c]
Cyclohexene		2.955	1.647	−3.601	−1.6	18.1	−3.6	−1.8	0.019	0.089	[c]
Cyclohexa-1,4-diene		0.779	1.984	−1.763	1.5	25.4	−0.8	2.9	0.014	0.084	[c]
Cyclohexa-1,3-diene		0.931	2.207	−2.138	3.2	30.0	0.8	7.3	0.031	0.078	[c]

[a] Reprinted, with permission, from Ref. [31]; Copyright 2005, American Institute of Physics.
[b] PDI cannot be computed for non 6-MRs.
[c] Nonplanar molecules that prevent easy and exact σ–π separation.

Table 15.2 Pearson coefficient (r^2) for correlation between different aromaticity indexes for the series of molecules in Table 15.1 at the HF/6-31G(d) level of theory.

	HOMA	NICS(0)	$NICS(0)_{zz}$	NICS(1)	$NICS(1)_{zz}$	PDI	FLU	FLU_π
HOMA	1.00	0.30	0.33	0.43	0.41	0.68	0.90	0.78
NICS(0)		1.00	0.98	0.96	0.97	0.50	0.47	0.67
$NICS(0)_{zz}$			1.00	0.96	0.98	0.64	0.49	0.70
NICS(1)				1.00	0.99	0.55	0.60	0.72
$NICS(1)_{zz}$					1.00	0.60	0.58	0.72
PDI						1.00	0.79	0.53
FLU							1.00	0.89
FLU_π								1.00

15.5 Applications of QTAIM to Aromaticity Analysis

In this section we present some applications of the QTAIM to the analysis of aromaticity in buckybowls, fullerenes, substituted and unsubstituted PAHs, and in the simplest Diels–Alder reaction.

15.5.1 Aromaticity of Buckybowls and Fullerenes

Different studies have attributed ambiguous aromatic character to fullerenes [36], which are regarded as aromatic by some criteria and as nonaromatic by others [37–39]. For example, their magnetic properties are indicative of extensive cyclic π-electron delocalization and substantial ring currents; evidence against their aromaticity, however, is that they are very reactive. We have analyzed the local aromaticity of fullerenes and buckybowls [40, 41], which also have fullerene-like physicochemical properties, to clarify their aromatic character. Table 15.3 lists the HF/6-31G(d)//AM1 values of the NICS, PDI, and HOMA measures of aromaticity, with the average pyramidalization angles, for a series of planar and bowl-shaped PAHs and fullerenes, i.e. from benzene to buckminsterfullerene (C_{60}), represented in Scheme 15.2.

The three local aromaticity criteria almost give the same trend for the different rings of the PAHs studied. Clear aromatic character is assigned to the hexagonal rings of benzene, naphthalene, and $C_{20}H_{10}$ and to the outer 6-MRs of $C_{26}H_{12}$ (C) and $C_{30}H_{12}$ (C and D), whereas the inner 6-MRs of $C_{26}H_{12}$ (A), $C_{30}H_{12}$ (A), and C_{60} are found to be moderately aromatic. In contrast with the significant local aromaticity of 6-MRs, the 5-MRs have antiaromatic character. It is also worth noticing some convergence in the local aromaticity of the inner 6-MRs when go-

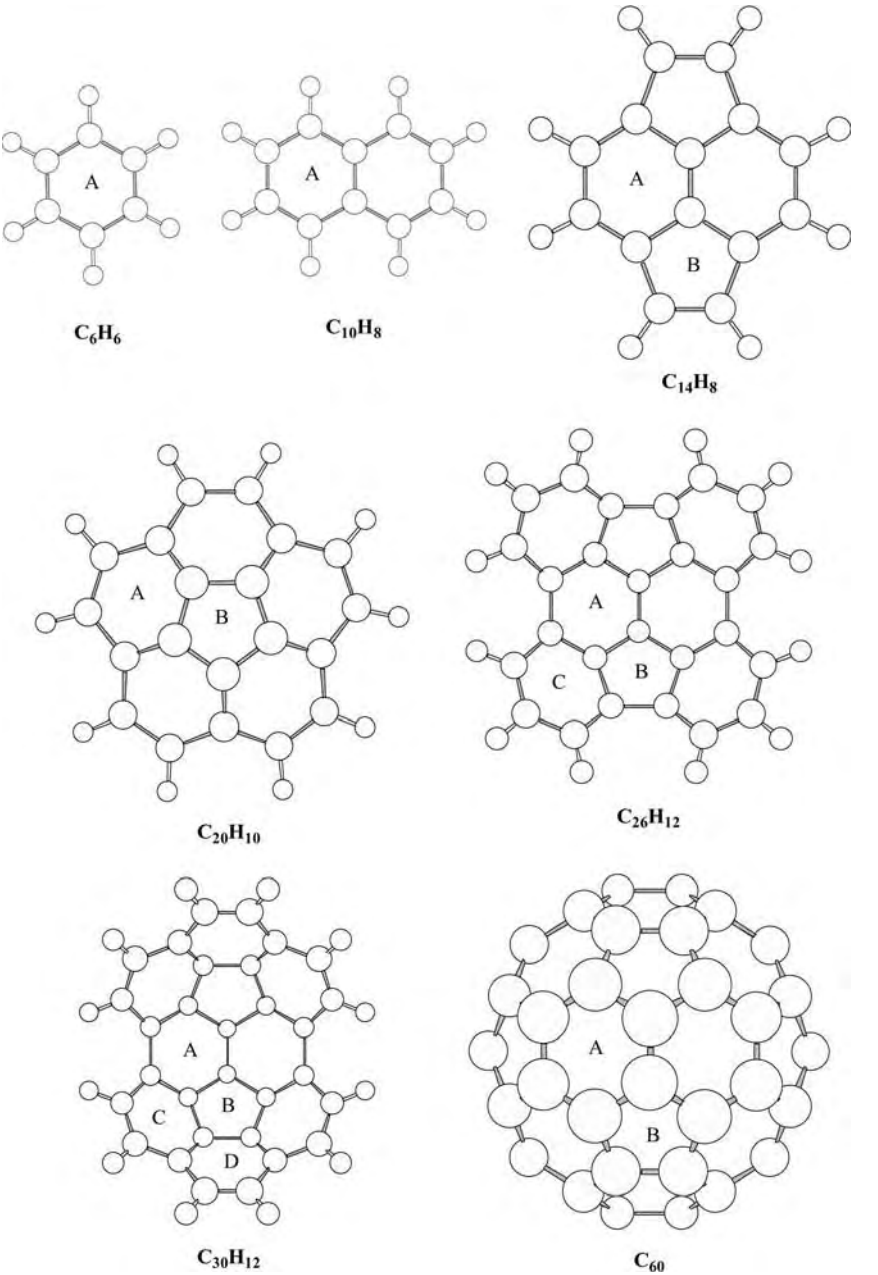
Table 15.3 HF/6-31G(d)//AM1 calculated values of NICS (ppm), HOMA, para-delocalization (PDI) (electrons) indices, and average pyramidalization angles for the carbon atoms present in a given ring (Pyr, in degrees) for a series of aromatic compounds.[a]

Molecule	Ring	NICS	HOMA	PDI	Pyr
C_6H_6	6A	−11.7	0.987	0.101	0.0
$C_{10}H_8$	6A	−11.3	0.807	0.074	0.0
$C_{14}H_8$	6A	−2.7	0.603	0.067	0.0
	5B	13.1	−0.205		0.0
$C_{20}H_{10}$	6A	−8.6	0.652	0.058	4.6
	5B	7.6	0.357		9.1
$C_{26}H_{12}$	6A	−5.6	0.474	0.037	6.9
	5B	3.9	−0.142		6.3
	6C	−10.0	0.746	0.078	2.8
$C_{30}H_{12}$	6A	−6.5	0.390	0.043	9.2
	5B	6.8	0.113		10.1
	6C	−9.4	0.652	0.061	5.1
	6D	−8.1	0.614	0.057	4.6
C_{60}	6A	−6.8	0.256	0.046	11.6
	5B	6.3	−0.485		11.6
C_{70}	5A	2.8	−0.481		11.9
	6B	−11.5	0.294	0.046	11.8
	5C	−1.3	−0.301		11.0
	6D	−8.8	0.141	0.028	10.1
	6E	−17.3	0.697	0.059	9.6

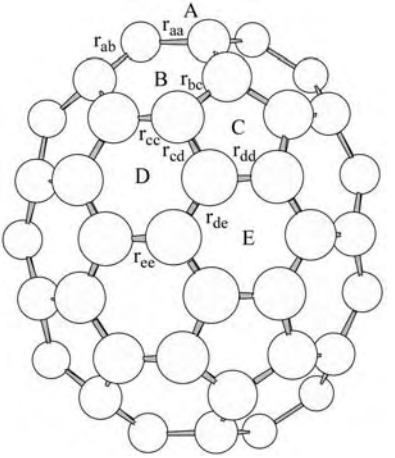
[a] Adapted, with permission, from Ref. [40]; copyright 2003, Wiley–VCH.

ing from the most aromatic benzene to the partially aromatic C_{60}. Unexpectedly for the bowl-shaped PAHs, the most pyramidalized outer rings have the largest local aromaticities, even though this trend agrees with the Clar model of aromaticity, which attributes more aromaticity to outer rings [42].

Another interesting fullerene is C_{70} [40], the structure of which arises from the insertion of an equatorial belt of five 6-MRs into C_{60} (Scheme 15.3), causing it to have different reactivity and local aromaticity [43]. Experimentally, the pole is less planar than the equatorial belt, which should imply lower aromaticity of the former. The HF/6-31G(d)//AM1 aromaticity criteria in Table 15.3 confirm that ring E, as expected, has the largest aromaticity, followed by ring B, located at the pole and, unexpectedly, by ring D, even though this is located in the equatorial belt. Compared with C_{60}, rings B and E of C_{70} are more aromatic than 6-MRs of C_{60}, in line with the accepted greater aromaticity of C_{70} [37]; despite this is more reactive [38]. NICS values of C_{70} are, in contrast, surprisingly high compared with those for C_{60}, even though they have similar geometric environments and pyramidalization angles. It must be noticed that experimental chemical shifts in



Scheme 15.2 Reprinted, with permission, from Ref. [40]; copyright 2003, Wiley–VCH.



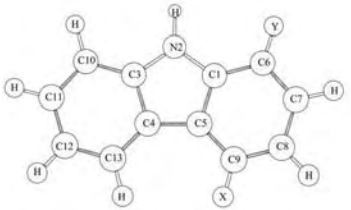
Scheme 15.3 Reprinted, with permission, from Ref. [40]; copyright 2003, Wiley–VCH.

^{3}He@C_{70} and ^{3}He@C_{60} follow the same trend as calculated NICS for C_{70} and C_{60} [44].

15.5.2 Effect of Substituents on Aromaticity

Benzene is regarded as the archetype of aromaticity, fulfilling all the criteria attributed to this property [2]. This molecule has been used as the reference for the proposal of quantitative descriptors of substituent effects, that is, the Hammett substituent constants [45]. We have tried to establish a relationship between the substituent effect and the aromaticity of a series of monosubstituted derivatives of benzene [46]. Table 15.4 contains the HF/6-31+G(d)//B3LYP/6-311+G(d,p) NICS, the B3LYP/6-311+G(d,p) HOMA, and the B3LYP/6-311G(d,p)//B3LYP/6-311+G(d,p) PDI aromaticity measures, with different substituent constants (explained elsewhere [46, 47]). It is apparent that, although the nature of the substituents varies substantially along the series (σ_p varying from -0.66 for a strongly electron-donating NH_2 substituent to 1.91 for a strongly electron-accepting NN^+ substituent), no significantly large changes of aromaticity are observed. This proves the high resistance of the π-electron structure of benzene, in agreement with its preference for substitution rather than addition reactions. In addition, PDI is the only index that gives a direct correlation between aromaticity and the substituent constants (Fig. 15.1), thus proving to be a good descriptor of changes of π-electron delocalization in substituted benzenes [46].

In another study the substituent effect was analyzed for a series of carbazole derivatives (Scheme 15.4) and an attempt was made to predict the reactivity of these systems quantitatively as a function of the substituent by measuring differ-



	1	2	3	4	5	6	7	8	9
X	H	-	CH_3	OH	Br	CN	$COCH_3$	COOH	CH_3
Y	H	H	H	H	H	H	H	H	$COCH_3$

Scheme 15.4 Reprinted, with permission, from Ref. [48]; copyright 2004, Royal Society of Chemistry.

ent local aromaticity criteria [48]. As is apparent from Fig. 15.2 for the substituted ring, the results for the three aromaticity criteria are scattered over a narrow range of values. As in the previous example, the π-electron structure of the aromatic ring is slightly affected by substituents. There is also clear divergence in

Table 15.4 GIAO/HF/6-31+G(d) NICS, B3LYP/6-311+G(d,p) HOMA, and B3LYP/6-311G(d,p) PDI aromaticity indices, calculated at B3LYP/6-311+G(d,p) geometry, for differently substituted benzene (C_6H_5X) structures. Also listed are substituent constants σ^+, σ^-, σ_m, and σ_p and resonance constants R^+ and R^- (m and p refer to *meta* and *para* substitution and + and − indicate the ability of the substituent to effectively delocalize either a positive or negative charge).[a]

−X	NICS (ppm)	HOMA	PDI (electrons)	σ^+/σ^-	σ_m	σ_p	R^+/R^-
$-NN^+$	−10.6	0.96	0.080	3.43	1.76	1.91	1.85
−NO	−9.8	0.98	0.091	1.63	0.62	0.91	1.14
$-NO_2$	−10.9	0.99	0.096	1.27	0.71	0.78	0.62
−CN	−10.3	0.98	0.096	1	0.56	0.66	0.49
−COCl	−9.9	0.98	0.095	1.24	0.51	0.61	0.78
$-COCH_3$	−9.7	0.98	0.097	0.84	0.38	0.5	0.51
$-COOCH_3$	−9.8	0.98	0.097	0.75	0.37	0.45	0.14
−COOH	−9.7	0.98	0.097	0.77	0.37	0.45	0.43
−CHO	−9.6	0.97	0.095	1.03	0.35	0.42	0.70
$-CONH_2$	−9.9	0.98	0.098	0.61	0.28	0.36	0.35
−CCH	−10.1	0.97	0.096	0.53	0.21	0.23	0.31
−Cl	−10.7	0.99	0.099	0.19	0.37	0.23	−0.31
−F	−11.7	0.99	0.098	−0.03	0.34	0.06	−0.52
−H	−9.7	0.99	0.103	0	0	0	0
−Ph	−9.3	0.98	0.098	−0.18	0.06	−0.01	−0.30
$-CH_3$	−9.7	0.98	0.100	−0.31	−0.07	−0.17	−0.32
$-OCH_3$	−10.8	0.98	0.094	−0.78	0.12	−0.27	−1.07
$-NH_2$	−9.8	0.98	0.093	−1.3	−0.16	−0.66	−1.38
−OH	−10.8	0.99	0.095	−0.92	0.12	−0.37	−1.25

[a] Reprinted, with permission, from Ref. [46]; copyright 2004, American Chemical Society.

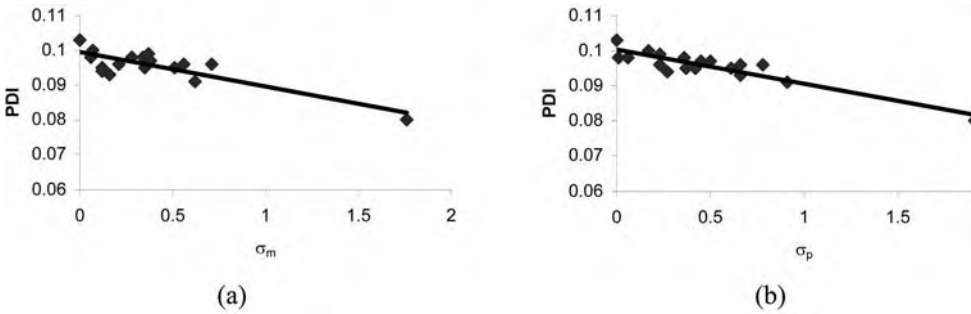


Fig. 15.1 Plot of PDI (electrons) against σ_m (a) and σ_p (b). The correlation coefficients are −0.83 and −0.91, respectively. Reprinted, with permission, from Ref. [46]; copyright 2004, American Chemical Society.

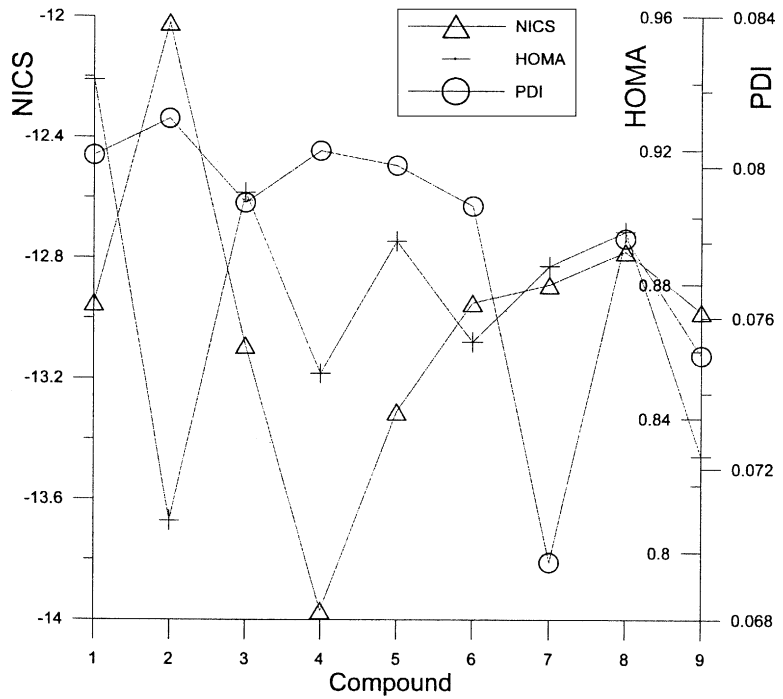


Fig. 15.2 Comparative plot of HOMA, NICS (ppm), and PDI (electrons) for the substituted ring (Subs) for the series of carbazole derivatives studied. The x-plot enumeration can be found in Scheme 15.4. Adapted, with permission, from Ref. [48]; copyright 2004, Royal Society of Chemistry.

Table 15.5 NICS (ppm), PDI (electrons), HOMA, and FLU measures of the aromaticity of the five and six-membered rings of guanine (G) and six-membered ring of cytosine (C) computed by the B3LYP method.[a]

System	NICS (ppm)			PDI (electrons)		HOMA			FLU		
	G-5	G-6	C-6	G-6	C-6	G-5	G-6	C-6	G-5	G-6	C-6
GC	−11.94	−4.10	−1.86	0.036	0.040	0.848	0.795	0.703	0.025	0.033	0.035
$[GC]^+$	−5.41	−0.31	−2.49	0.023	0.042	0.829	0.550	0.773	0.028	0.048	0.031
Ca^{2+}-GC	−10.67	−4.76	−2.53	0.044	0.045	0.843	0.886	0.797	0.023	0.024	0.029
Cu^{+}-GC	−10.64	−4.59	−2.25	0.040	0.043	0.869	0.898	0.761	0.021	0.027	0.032
Cu^{2+}-GC	−7.37	−2.00	−3.07	0.022	0.040	0.915	0.760	0.822	0.033	0.058	0.039

[a] Adapted, with permission, from Ref. [49]; copyright 2005, Taylor and Francis. Details of the basis set used can be found in Ref. [49].

the ordering of the different systems by local aromaticity values given by the B3LYP/6-31++G(d,p) values of NICS, HOMA, and PDI, at variance with previous analyses of aromaticity which resulted in relatively good agreement among the different aromaticity criteria. One must be very cautious with the results, because it is not being possible to give a definite answer about the relative aromaticity of these rings. Finally, it is worth saying that the different trends found by use of these criteria are not completely unexpected, because we are considering descriptors based on different physical properties.

The last study on this subject has consisted in the analysis of how metal cations and ionization affect the aromaticity of the Watson–Crick guanine–cytosine base pair (GC) [49]. Table 15.5 contains the B3LYP aromaticity results (HOMA, PDI, NICS, and FLU) for the 6-MRs and 5-MRs studied (Fig. 15.3). H-bond formation in GC implies some loss of π-charge in N_1 and a gain in O_6, thus increasing the relevance of the resonance structure **2** (Fig. 15.3), which favors intensification of

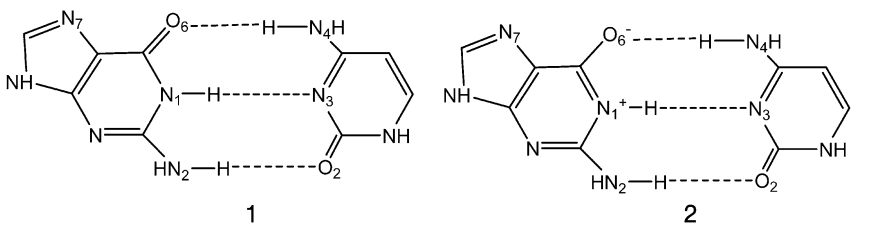
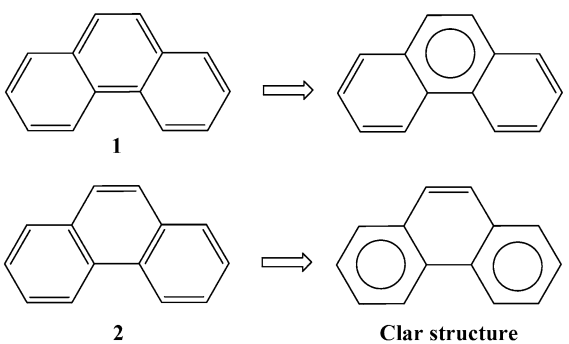


Fig. 15.3 Schematic representation of charge transfer in the GC base pair. G is guanine and C is cytosine. Reprinted, with permission, from Ref. [49]; copyright 2005, Taylor and Francis.

the aromatic character of the guanine 6-MR. The increased aromaticity of the guanine and cytosine 6-MRs which results from interactions with Cu^{+} and Ca^{2+} is also attributed to the strengthening of hydrogen-bonding in the GC pair, which stabilizes the charge-separation resonance structure **2** [50]. This effect is stronger for the divalent Ca^{2+} metal cation than for the monovalent Cu^{+}. The observed reduction of the aromaticity of the 5-MRs and 6-MRs of guanine because of ionization or interaction with Cu^{2+} is caused by the oxidation process that removes one π electron, disrupting the π-electron distribution.

15.5.3
Assessment of Clar's Aromatic π-Sextet Rule

The introduction of Hückel's $4n + 2$ rule enabled better comprehension of aromaticity [51], although strictly it could only be applied to monocyclic conjugated systems. This was solved later by Clar's model of the extra stability of $6n$ π-electron benzenoid species [42, 52]. According to Clar's rule, the Kekulé resonance structure with the largest number of disjoint aromatic π-sextets, i.e. benzene-like moieties, is the most important for the characterization of the properties of polycyclic aromatic hydrocarbons (PAHs). Aromatic π-sextets are defined as six π-electrons localized in a single benzene-like ring separated from adjacent rings by formal C–C single bonds. For example, application of this rule to phenanthrene reveals that resonance structure **2** (Scheme 15.5) is more important than resonance structure **1**; which is translated into major local aromaticity of the outer rather than inner rings. The Clar structure of a given PAH is the resonance structure with the maximum number of isolated and localized aromatic π-sextets, with a minimum number of localized double bonds. A PAH with more π-sextets is kinetically more stable than its isomer with fewer. In addition, π-sextets are regarded as the most aromatic rings of PAHs. Some PAHs (e.g. phenanthrene) have a single Clar structure whereas others have several. For these latter, Clar's rule cannot differentiate which of the resonance structures is mainly responsible



Scheme 15.5 Reprinted, with permission, from Ref. [53]; copyright 2005, Wiley–VCH.

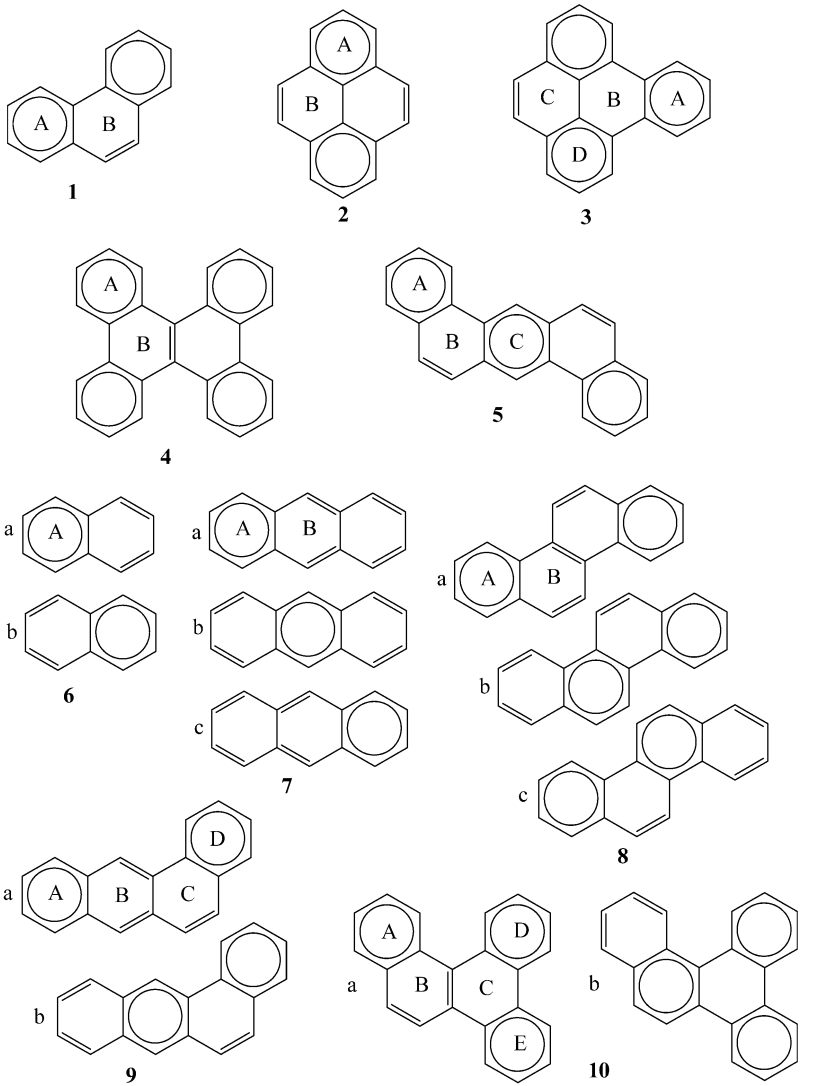
Table 15.6 PDI (in electrons), HOMA, and NICS (in ppm) values for the PAHs studied. The numbering is given in Scheme 15.6.[a]

	Ring	Molecule									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
PDI	A	0.080	0.069	0.086	0.084	0.083	0.076	0.066	0.080	0.069	0.079
	B	0.047	0.043	0.026	0.034	0.041		0.066	0.053	0.066	0.057
	C			0.044		0.068				0.038	0.031
	D			0.073						0.084	0.085
	E										0.085
HOMA	A	0.856	0.834	0.889	0.811	0.872	0.769	0.619	0.829	0.697	0.749
	B	0.435	0.553	−0.030	0.383	0.356		0.696	0.542	0.730	0.305
	C			0.518		0.788				0.266	−0.097
	D			0.838						0.883	0.820
	E										0.883
NICS	A	−10.06	−12.74	−8.63	−9.44	−9.93	−9.98	−8.84	−9.94	−9.30	−10.19
	B	−6.82	−5.07	−1.18	−4.13	−5.38		−12.60	−7.69	−11.69	−7.68
	C			−5.47		−11.27				−4.58	−3.91
	D			−11.58						−9.81	−9.55
	E										−8.99

[a] Adapted, with permission, from Ref. [53]; copyright 2005, Wiley–VCH.

for the aromaticity of the system. In this study we investigate whether three local aromaticity criteria, PDI, HOMA, and NICS, give results consistent with Clar's original qualitative π-sextet rule [53].

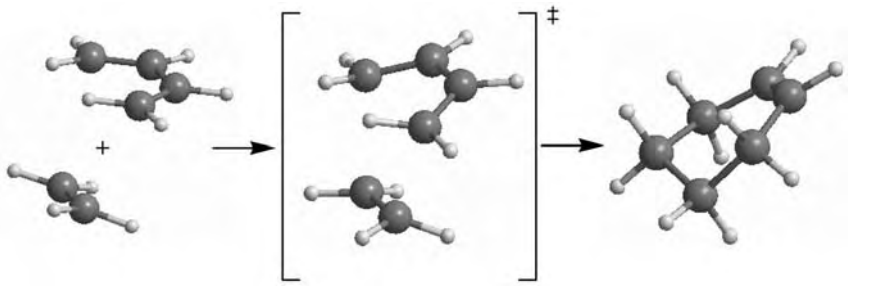
The PAHs studied are depicted in Scheme 15.6. PAHs **1–5** have a single Clar structure and **6–10** have several Clar valence structures, also represented. The corresponding local aromaticity values, calculated at the B3LYP/6-31G(d) level, can be found in Table 15.6. First, for systems with a single Clar structure (**1–5**), π-sextet rings have higher PDI values, higher HOMA, and more negative NICS than non-π-sextet rings. Hence, all three aromaticity criteria used agree perfectly with the qualitative description given by Clar's rule. Second, for systems with several Clar structures, it is apparent that the overall aromaticity of the system given by PDI and HOMA agrees with the superimposition of all possible Clar structures. For example, for **7**, PDI and HOMA attribute very similar aromaticity to rings A and B, which proves the non-localizability of the π-sextet. NICS, however, attributes much more aromatic character to ring B than ring A, although it is claimed this is because of overestimation by NICS of the local aromaticity of the inner rings of PAHs [54].



Scheme 15.6 Adapted, with permission, from Ref. [53]; copyright 2005, Wiley–VCH.

15.5.4
Aromaticity Along the Diels–Alder Reaction. The Failure of Some Aromaticity Indexes

This work [33] analyzes the aromaticity along the Diels–Alder reaction between 1,3-butadiene and ethane to yield cyclohexene (Scheme 15.7) [55, 56], which is often taken as a prototypical pericyclic concerted reaction. This reaction is characterized by an aromatic transition state (TS) [56, 57], thus along the reaction path a



Scheme 15.7 Reprinted, with permission, from Ref. [33]; copyright 2005, Elsevier.

peak of aromaticity around the TS is expected. The trends of aromaticity along the path for the different criteria applied, at the B3LYP/6-31G(d) level, can be found in Fig. 15.4. Only the magnetic NICS(1) and the electronic PDI criteria find the most aromatic point along the reaction path around the TS of the reaction. In contrast, both geometric HOMA and electronic FLU regard cyclohexene as the most aromatic species in this reaction. This latter trend is also given by the RSS (root summed squares) of the best fitted plane for atoms in the ring, which is an unambiguous measure of molecular planarity [33]. It shows the product as the most planar species, thus in principle implying greater π-electron delocalization. The failure of RSS proves that the flatter structure is not necessarily the more aromatic. On the other hand, HOMA and FLU measure variances of structural and electronic patterns around the ring, and might fail if they are not applied to stable species, for example in a reaction with major structural and electronic changes. The failure of some indices to detect the aromaticity of the TS in the simplest Diels–Alder cycloaddition thus reinforces the idea of the multidimensional character of aromaticity [8] and the need to use several criteria to quantify it.

15.6 Conclusions

A key aspect of aromatic compounds is the π-electron delocalization (and σ and even δ-electron delocalization in all-metal and inorganic aromatic species) present in these molecules. In this chapter we have defined three new aromaticity indexes founded on evaluation of electron delocalization in the framework of the QTAIM, i.e. the para-delocalization (PDI), aromatic fluctuation (FLU), and FLU_π indexes. We have shown that theoretical studies of electron delocalization using QTAIM-based tools have significantly improved our understanding of aromaticity in fullerenes, substituted benzene derivatives, polycyclic aromatic hydrocarbons, and chemical reactivity. The lack of a universally accepted measure of aromaticity, its multidimensional character, and the limitations of almost all descriptors of aromaticity stress the need for new aromaticity criteria in addition to those defined

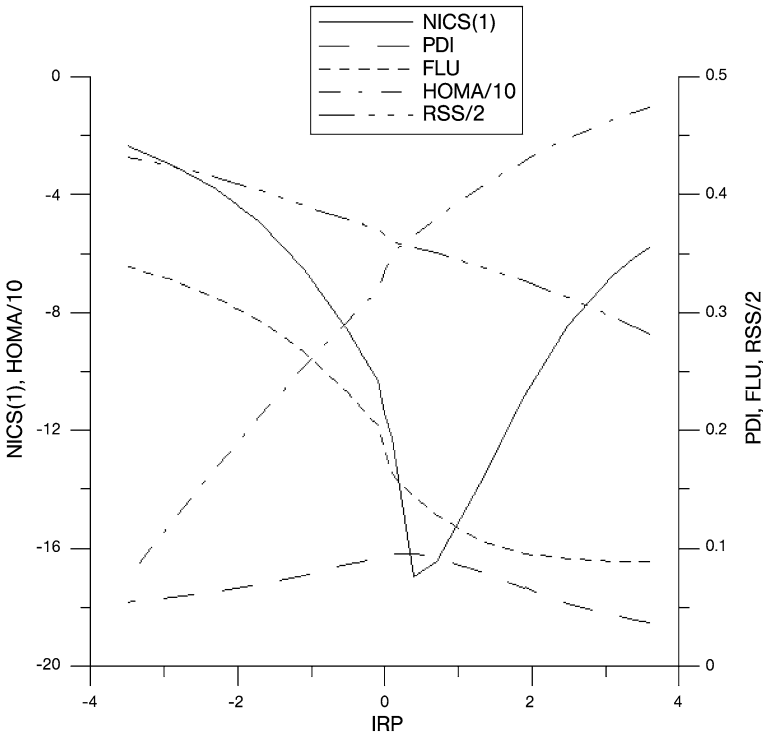


Fig. 15.4 Plot of NICS(1) (ppm), PDI (electrons), FLU, HOMA (values divided by 10), and RSS (values divided by 2) against the reaction coordinate (IRP in amu$^{1/2}$ bohr). Negative values of the IRP correspond to the reactants side of the reaction path, positive values to the product side, and IRP = 0.000 corresponds to the TS of the DA cycloaddition. Reprinted, with permission, from Ref. [33]; copyright 2005, Elsevier.

in this chapter. In a given study of a series of compounds, one can safely reach a definite conclusion about their aromaticity only when differently based indicators of aromaticity lead to the same results. For this reason, in our opinion, careful analysis of the aromaticity of a set of molecules must be performed using electronically based descriptors, for example the PDI or FLU indexes, and geometry-based indicators such as the HOMA index, magnetically based measures, for example NICS, and energetically based descriptors, for example ASEs.

Acknowledgments

The authors are grateful to Professor Dr Miquel Duran and Dr Xavier Fradera for helpful discussions. Financial help was obtained from the Spanish Ministerio de Educación y Ciencia (MEC) project No. CTQ2005-08797-C02-01/BQU and from

the Catalan Departament d'Universitats, Recerca i Societat de la Informació (DURSI) project No. 2005SGR-00238. E.M. and J.P. thank the MEC for doctoral fellowship no. AP2002-0581 and the DURSI for the postdoctoral fellowship 2005BE00282, respectively. Excellent service by the Centre de Supercomputació de Catalunya (CESCA) is gratefully acknowledged.

References

1 P. v. R. Schleyer, *Chem. Rev.* **2001**, *101*, 1115–1118; P. v. R. Schleyer, *Chem. Rev.* **2005**, *105*, 3433–3435; V. I. Minkin, M. N. Glukhovtsev, B. Y. Simkin, *Aromaticity and Antiaromaticity: Electronic and Structural Aspects*, John Wiley and Sons, New York, **1994**; P. J. Garrat, *Aromaticity*, John Wiley and Sons, New York, **1986**; A. T. Balaban, *Pure Appl. Chem.* **1980**, *52*, 1409–1429; D. Lloyd, *J. Chem. Inf. Comp. Sci.* **1996**, *36*, 442–447.

2 T. M. Krygowski, M. K. Cyrański, Z. Czarnocki, G. Häfelinger, A. R. Katritzky, *Tetrahedron* **2000**, *56*, 1783–1796.

3 X. Li, A. E. Kuznetsov, H.-F. Zhang, A. Boldyrev, L.-S. Wang, *Science* **2001**, *291*, 859–861; A. I. Boldyrev, L.-S. Wang, *Chem. Rev.* **2005**, *105*, 3716–3757; C. A. Tsipis, *Coord. Chem. Rev.* **2005**, *249*, 2740–2762.

4 Z. Chen, C. S. Wannere, C. Corminboeuf, R. Puchta, P. v. R. Schleyer, *Chem. Rev.* **2005**, *105*, 3842–3888.

5 P. Lazzeretti, in *Progress in Nuclear Magnetic Resonance Spectroscopy* (Ed.: J. W. Emsley, J. Feeney, and L. H. Sutcliffe), Elsevier, Amsterdam, **2000**, p. 1–88.

6 M. Faraday, *Phil. Trans. Roy. Soc. London* **1825**, 440.

7 J. Poater, M. Duran, M. Solà, B. Silvi, *Chem. Rev.* **2005**, *105*, 3911–3947.

8 A. R. Katritzky, M. Karelson, S. Sild, T. M. Krygowski, K. Jug, *J. Org. Chem.* **1998**, *63*, 5228–5231; A. R. Katritzky, K. Jug, D. C. Oniciu, *Chem. Rev.* **2001**, *101*, 1421–1449; T. M. Krygowski, M. K. Cyrański, *Chem. Rev.* **2001**, *101*, 1385–1419; M. K. Cyrański, T. M. Krygowski, A. R. Katritzky, P. v. R. Schleyer, *J. Org. Chem.* **2002**, *67*, 1333–1338.

9 K. Jug, A. M. Köster, *J. Phys. Org. Chem.* **1991**, *4*, 163–169.

10 J. Kruszewski, T. M. Krygowski, *Tetrahedron Lett.* **1972**, *13*, 3839–3842; T. M. Krygowski, *J. Chem. Inf. Comp. Sci.* **1993**, *33*, 70–78.

11 T. M. Krygowski, M. C. Cyrański, *Tetrahedron* **1996**, *52*, 10255–10264; T. M. Krygowski, M. K. Cyrański, *Tetrahedron* **1996**, *52*, 1713–1722.

12 P. v. R. Schleyer, C. Maerker, A. Dransfeld, H. Jiao, N. J. R. van Eikema Hommes, *J. Am. Chem. Soc.* **1996**, *118*, 6317–6318.

13 M. K. Cyrański, *Chem. Rev.* **2005**, *105*, 3773–3811.

14 F. De Proft, P. Geerlings, *Chem. Rev.* **2001**, *101*, 1451–1464.

15 R. F. W. Bader, *Acc. Chem. Res.* **1985**, *18*, 9–15; R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Clarendon, Oxford, **1990**; R. F. W. Bader, *Chem. Rev.* **1991**, *91*, 893–928.

16 E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry*, Academic, New York, **1976**; R. McWeeny, in *Methods of Molecular Quantum Mechanics*, Academic Press, London, **1992**; J. Cioslowski, *Many-Electron Densities and Reduced Density Matrices*, Kluwer Academic/Plenum Publishers, New York, **2000**.

17 K. Ruedenberg, *Rev. Mod. Phys.* **1962**, *34*, 326–376.

18 R. F. W. Bader, M. E. Stephens, *Chem. Phys. Lett.* **1974**, *26*, 445–449.

19 R. F. W. Bader, M. E. Stephens, *J. Am. Chem. Soc.* **1975**, *97*, 7391–7399.

20 M. A. Buijse, E. J. Baerends, in *Density Functional Theory of Molecules*,

Clusters and Solids (Ed.: D. E. Ellis), Kluwer, Dordrecht, **1995**, p. 1–46; E. J. Baerends, O. V. Gritsenko, *J. Phys. Chem. A* **1997**, *101*, 5383–5403; M. A. Buijse, E. J. Baerends, *Mol. Phys.* **2002**, *100*, 401–421.

21 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A* **1999**, *103*, 304–314.

22 R. F. W. Bader, A. Streitwieser, A. Neuhaus, K. E. Laidig, P. Speers, *J. Am. Chem. Soc.* **1996**, *118*, 4959–4965.

23 R. L. Fulton, S. T. Mixon, *J. Phys. Chem.* **1993**, *97*, 7530–7534.

24 E. Matito, M. Solà, P. Salvador, M. Duran, *Faraday Discuss.* **2006**, in press.

25 F. Sondheimer, *Pure Appl. Chem.* **1964**, *7*, 363–388.

26 P. v. R. Schleyer, H. Jiao, *Pure Appl. Chem.* **1996**, *68*, 209–218.

27 C. F. Matta, J. Hernández-Trujillo, *J. Phys. Chem. A* **2003**, *107*, 7496–7504; C. F. Matta, J. Hernández-Trujillo, *J. Phys. Chem. A* **2005**, *109*, 10798.

28 P. Bultinck, M. Rafat, R. Ponec, B. van Gheluwe, R. Carbó-Dorca, P. Popelier, *J. Phys. Chem. A* **2006**, *110*, 7642–7648.

29 J. Poater, X. Fradera, M. Duran, M. Solà, *Chem. Eur. J.* **2003**, *9*, 400–406.

30 C. W. Bird, *Tetrahedron* **1985**, *41*, 1409–1414.

31 E. Matito, M. Duran, M. Solà, *J. Chem. Phys.* **2005**, *122*, 014109.

32 J. M. Matxain, L. A. Eriksson, J. M. Mercero, X. López, J. M. Ugalde, J. Poater, E. Matito, M. Solà, **2006**, to be submitted.

33 E. Matito, J. Poater, M. Duran, M. Solà, *J. Mol. Struct. (Theochem)* **2005**, *727*, 165–171.

34 I. Mayer, P. Salvador, *Chem. Phys. Lett.* **2004**, *383*, 368–375.

35 E. Matito, J. Poater, M. Solà, M. Duran, P. Salvador, *J. Phys. Chem. A* **2005**, *109*, 9904–9910; E. Matito, P. Salvador, M. Duran, M. Solà, *J. Phys. Chem. A* **2006**, *110*, 5108–5113.

36 L. T. Scott, J. S. Siegel, *Tetrahedron* **2001**, *57*, ix.

37 M. Bühl, A. Hirsch, *Chem. Rev.* **2001**, *101*, 1153–1183.

38 R. C. Haddon, L. F. Schneemeyer, J. V. Waszczak, S. H. Glarum, R. Tycko, G. Dabbagh, A. R. Kortan, A. J. Muller, A. M. Mujse, M. J. Rosseinsky, S. M. Zahurak, A. V. Makhija, F. A. Thiel, K. Raghavachari, E. Cockayne, V. Elser, *Nature* **1991**, *350*, 46–47; T. M. Krygowski, A. Ciesielski, *J. Chem. Inf. Comp. Sci.* **1995**, *35*, 1001–1003.

39 P. Fowler, *Nature* **1991**, *350*, 20–21; R. C. Haddon, *Science* **1993**, *261*, 1545–1550.

40 J. Poater, X. Fradera, M. Duran, M. Solà, *Chem. Eur. J.* **2003**, *9*, 1113–1122.

41 J. Poater, X. Fradera, M. Duran, M. Solà, in *Fullerenes: The Exciting World of Nanocages and Nanotubes* (Ed.: P. V. Kamat, K. M. Kadish, and D. Guldi), The Electrochemical Society Inc., Pennington, **2002**, p. 707–719.

42 E. Clar, *The Aromatic Sextet*, Wiley, New York, **1972**.

43 R. Taylor, J. P. Hare, A. K. Abdul-Sada, H. W. Kroto, *J. Chem. Soc., Chem. Commun.* **1990**, 1423–1424; J. Mestres, M. Duran, M. Solà, *J. Phys. Chem.* **1996**, *100*, 7449–7454; T. Akasaka, E. Mitsuhida, W. Ando, K. Kobayashi, S. Nagase, *J. Am. Chem. Soc.* **1994**, *116*, 2627–2628; A. L. Balch, J. W. Lee, M. M. Olmstead, *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 1356–1358; C. Bellavia-Lund, F. Wudl, *J. Am. Chem. Soc.* **1997**, *119*, 9937–9937; M. S. Meier, G. W. Wang, R. C. Haddon, C. P. Brock, M. A. Lloyd, J. P. Selegue, *J. Am. Chem. Soc.* **1998**, *120*, 2337–2342; C. Thilgen, A. Herrmann, F. Diederich, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2269–2280.

44 Z. Chen, R. B. King, *Chem. Rev.* **2005**, *105*, 3613–3642.

45 L. P. Hammett, *Chem. Rev.* **1935**, *17*, 125–136.

46 T. M. Krygowski, K. Ejsmont, B. T. Stepién, M. K. Cyrański, J. Poater, M. Solà, *J. Org. Chem.* **2004**, *69*, 6634–6640.

47 C. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165–195.

48 J. Poater, I. García-Cruz, F. Illas, M. Solà, *Phys. Chem. Chem. Phys.* **2004**, *6*, 314–318.

49 J. Poater, M. Sodupe, J. Bertran, M. Solà, *Mol. Phys.* **2005**, *103*, 163–173.

50 M. K. Cyrański, M. Gilski, M. Jaskolski, T. M. Krygowski, *J. Org. Chem.* **2003**, *68*, 8607–8613; V. G. S. Box, F. Jean-Mary, *J. Mol. Mod.* **2001**, *7*, 334.

51 E. Hückel, *Z. Physik* **1931**, *70*, 104–186; E. Hückel, *Z. Physik* **1931**, *72*, 310–337; E. Hückel, *Z. Physik* **1932**, *76*, 628–648; E. Hückel, *Z. Elektrochemie* **1937**, *43*, 752–788, 827–849.

52 M. Randić, *Chem. Rev.* **2003**, *103*, 3449–3605.

53 G. Portella, J. Poater, M. Solá, *J. Phys. Org. Chem.* **2005**, *18*, 785–791.

54 G. Portella, J. Poater, J. M. Bofill, P. Alemany, M. Solá, *J. Org. Chem.* **2005**, *70*, 2509–2521.

55 K. N. Houk, J. Gonzalez, Y. Li, *Acc. Chem. Res.* **1995**, *28*, 81–90; R. Herges, H. Jiao, P. v. R. Schleyer, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1376.

56 H. Jiao, P. v. R. Schleyer, *J. Phys. Org. Chem.* **1998**, *11*, 655–662.

57 M. Manoharan, F. De Proft, P. Geerlings, *J. Org. Chem.* **2000**, *65*, 7971–7976; M. Manoharan, F. De Proft, P. Geerlings, *J. Org. Chem.* **2000**, *65*, 6132–6137; M. Manoharan, F. De Proft, P. Geerlings, *J. Chem. Soc., Perkin Trans. 2* **2000**, 1767–1773.

16
Topological Properties of the Electron Distribution in Hydrogen-bonded Systems

Ignasi Mata, Ibon Alkorta, Enrique Espinosa, Elies Molins, and José Elguero

16.1
Introduction

The hydrogen bond (HB) is the most important weak interaction found in nature. It is responsible for the three dimensional shape of biopolymers (proteins and nucleic acids) and for the structure of water, in both the liquid and solid phases. Life processes extensively use the making and breaking of HBs as part of concatenated reactions involving huge amounts of biomolecules. In addition, it has profound implications in the mode of action of drugs and in molecular packing, recognition, and crystal engineering [1].

In the literature, the moieties involved in HB interaction are usually identified as hydrogen donor (or electron acceptor) and hydrogen acceptor (or electron donor). In this chapter the donor (D) and acceptor (A) terms are schematically represented as D–H···A.

The H···A interaction distance d(H···A) varies from 1.2 to 2.5 Å (or up to 3.0 Å, depending of the criteria used). This is not true for covalent bonds, for which the range is much smaller. Dependence of hydrogen bond properties on the internuclear distance can, therefore, be clearly observed for hydrogen bonds (HBs), as shown by the D–H bond distance, which seems to depend on d(H···A). The mutual dependence of bond distances on both sides of the hydrogen atom can easily be understood in terms of the bond-order model proposed by Pauling, which assumes a total valence equal to 1 for the hydrogen atom involved in the HB interaction [2–5].

Similar dependencies have been found for the electron-density properties from application of QTAIM methodology to the hydrogen bond. Topological analysis of $\rho(\mathbf{r})$ was initially used to identify the presence of HB interactions. Thus, characterization of C–H···O hydrogen bonds has been used to generalize a set of criteria to establish the presence of hydrogen-bonding interactions based on the QTAIM theory [6]. These criteria have been applied to the study of other hydrogen bonds apart from C–H···O interactions, and a further extension has been

developed for dihydrogen bonds (D–H···H–M) [7]. More recently, topological analysis has revealed the dependence of electron density properties on H···A distance and the interdependences between these properties.

This chapter covers theoretical and experimental studies that analyze the dependence of the topological and energy properties of the electron distribution in HB systems either on the H···A distance or on other properties of $\rho(\mathbf{r})$. As a particular point of interest, Section 16.6 discusses attempts which have been made to analyze the $\rho(\mathbf{r})$ properties in a comprehensive range of interatomic distances, from covalent to weak HB interactions. In general, the theoretical studies considered here have been undertaken in two different manners. The first treats the complexes by following a reaction coordinate, where the HB distance between the two D–H and A moieties involved is artificially modified. The second approach assembles information from complexes that have different internuclear H···A distances at their equilibrium geometries. The former method has the advantage of reducing secondary effects that arise in the latter as a result of the different electronegativities of the D and Y groups in the D–H···A–Y systems. It should be noted, however, that, in the former instance, the observed dependencies do not necessarily follow the same quantitative behavior as systems in equilibrium. The second method can, moreover, be used with experimental electron densities in crystals. Occasionally both approaches have been combined in the same study, filling the gaps in the data for complexes in equilibrium with calculated magnitudes that have been obtained by making use of reaction coordinates.

16.2
Topological Properties of the Hydrogen Bond

16.2.1
Topological Properties at the Bond Critical Point (BCP)

Within the QTAIM theory a bond between two atoms is characterized by a line of maximum electron density (the bond path) that connects both nuclei and intersects the zero-flux surface of the electron density gradient field $\nabla\rho(\mathbf{r})$ at a topological $(3, -1)$ point, called the bond critical point (BCP).

Some particular properties of the electron distribution at the BCP are:

1. ρ_b is the smallest value of the electron density along the bond path; and
2. it corresponds to the maximum of $\rho(\mathbf{r})$ at the interatomic surface (IAS) between both atoms.

The uniqueness of this point has been used to define some characteristics of the whole bond (for example bond distance, degree of covalence, interaction energy, etc.) in terms of a number of $\rho(\mathbf{r})$ properties at BCP. The most studied topo-

logical properties at the BCP are the electron density (ρ_b), its Laplacian ($\nabla^2\rho_b$) and the eigenvalues of the Hessian matrix (λ_1, λ_2 and λ_3), the latter indicating the three main curvatures of $\rho(\mathbf{r})$ at the BCP. If $\rho(\mathbf{r})$ at the BCP has a saddle distribution, two of the eigenvalues have negative values and correspond to the curvatures that are perpendicular to the bond path and the third is positive and represents the curvature of the $\rho(\mathbf{r})$ distribution along the bond path. By convention, the negative curvatures are λ_1 and λ_2 ($\lambda_1 < \lambda_2 < 0$), and the positive one is λ_3. The Laplacian of the electron density $\nabla^2\rho(\mathbf{r})$, which is defined as the sum of the three eigenvalues of the Hessian, λ_1, λ_2 and λ_3, provides information about either the charge concentration ($\nabla^2\rho(\mathbf{r}) < 0$) or the charge depletion ($\nabla^2\rho(\mathbf{r}) > 0$) of the electron distribution. For hydrogen bonds, ρ_b is usually small and $\nabla^2\rho_b > 0$, both being characteristic magnitudes of closed-shell interactions. (See also Chapter 1 of this book).

The first publications in which topological analysis of $\rho(\mathbf{r})$ was used to characterize hydrogen-bonded systems included theoretical calculations [HF/6-31G(d,p)] of the complexes formed between nitrile derivatives and hydrogen halides (HF and HCl) [8, 9]. In these studies the ρ_b magnitudes calculated for each hydrogen halide were linearly correlated with the H···A distance. A similar linear correlation was also observed for $\nabla^2\rho_b$ [9]. The small range of HB distances (0.28 and 0.36 Å for XCN···HF and XCN···HCl complexes, respectively) induced researchers to postulate linear relationships rather than other kinds of dependence. Similar conclusions were also reached by other authors when they considered a small range of distances. When the range expanded, however, the curvature of the ρ_b and $\nabla^2\rho_b$ data distributions became evident, as shown by a theoretical study of carbenes and silylenes as HB acceptors [10] and experimental analysis of H···O interactions [11].

The first study describing the nonlinear behavior of the ρ_b magnitude with the bond distance does not correspond to the analysis of a weak interaction but to the S–S bond, which can expand by 0.5 Å ($1.70 < d(\text{S–S}) < 2.25$ Å) [12]. These results clearly showed that a double-logarithm function, which can be expressed as an equivalent power function, enables better fitting of the data distribution than a linear regression (Fig. 16.1). In the same way, the experimentally derived ρ_b magnitudes of the CO bonds in citrinin were better fitted by a simple exponential function than by a linear function [13].

In experimental analysis of X–H···O (X = C, N, O) hydrogen bond interactions in crystals [11], ρ_b, $\nabla^2\rho_b$, and the curvatures were exponentially dependent on intermolecular distance (Fig. 16.2). For the same set of HB interactions, an exponential relationship was also found between the sum of the negative curvatures $\lambda_1 + \lambda_2$ and the positive curvature λ_3 [14]. These dependencies have been also derived in theoretical studies of H···F hydrogen-bonding interactions [15]. It is particularly worthy of note that the fitting data found for ρ_b in experimental H···O and in the theoretically calculated pure closed-shell H···F interactions $\{\rho_b = 65(27) \times \exp[-3.2(2)d(\text{H}\cdots\text{O})]\}$ and $\{\rho_b = 63(10) \times \exp[-3.552(7)d(\text{H}\cdots\text{F})]\}$ are very similar.

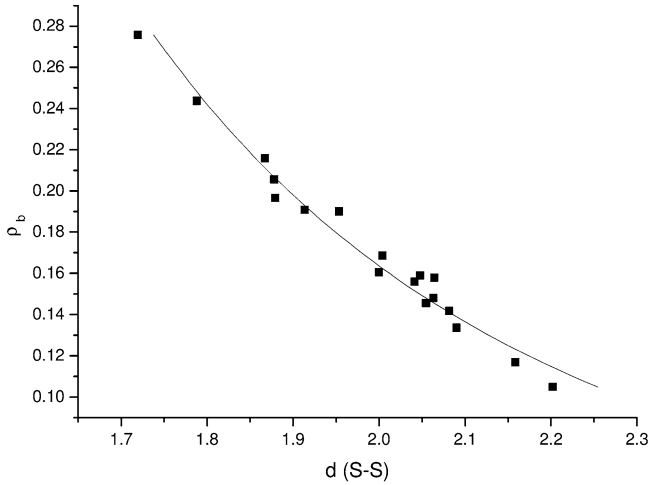


Fig. 16.1 Relationship between d(S–S) distance (Å) and ρ_b (a.u.). The fitted curve corresponds to $d(\text{S–S}) = 1.229\rho^{-0.269}$. (Values taken from Ref. [12]).

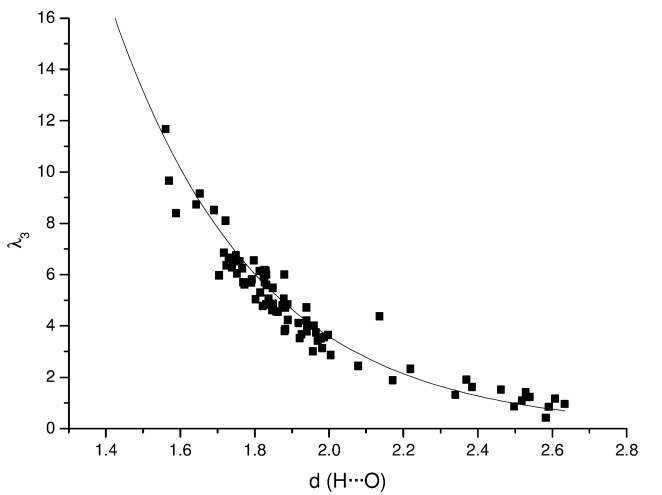


Fig. 16.2 Relationship between λ_3 (e Å^{-5}) and H···O distance (Å) data for experimental X–H···O (X = C, N, O) hydrogen-bonding interactions in crystals. (Values taken from Ref. [14]). The equation of the fitted curve is: $\lambda_3 = 0.65(13) \times 10^3 \exp[-2.6(1)d(\text{H}\cdots\text{O})]$.

16.2.2 Integrated Properties

Within a molecule, each atomic basin Ω is defined as the space region covered by all the gradient paths of $\rho(\mathbf{r})$ ending at its nuclear position. These atomic regions are delimited by zero-flux surfaces of $\nabla\rho(\mathbf{r})$ which do not belong to the atoms and represent their topological borders. From QTAIM methodology, atomic basins can be regarded as quantum subsystems of the molecular quantum system. Thus, the magnitude of a given molecular property P can be obtained by using the integration within the basins Ω, which provides the atomic contributions P_Ω to the molecular property P ($P = \Sigma_\Omega P_\Omega$). Several integrated properties have been analyzed in theoretical studies of HBs, for example charge, energy, volume, and atomic dipole polarizability. On HB formation, the hydrogen atom involved in the interaction becomes more positively charged, energetically destabilized, and its volume and dipole moment decrease. These changes were first observed for weak C–H···O hydrogen bonds by comparing the integrated properties of the complexes with those of the free monomers [6]. They have been included as four of the eight criteria used to establish hydrogen-bonding interactions from topological analysis of $\rho(\mathbf{r})$. When an HB is formed, electrons are typically promoted from the lone pair of the acceptor to the σ^* antibonding orbital of D–H (except for blue shift hydrogen bonds, in which other orbitals of the donor molecule are involved [16–18]), weakening the covalent bond and producing changes in the energy, charge, and dipole moment of the hydrogen atomic basin that are intrinsically related to each other. Indeed, as shown for hydrogen bonds in dimers of tetrahydroimidazo[4,5-*d*]imidazole derivatives [19], there is a clear correlation between the variations of the net charge and the energy of the hydrogen atom on HB formation (Fig. 16.3).

The shrinking of the hydrogen atom volume is explained by the proximity of the HB acceptor and the concomitant overlap of their electron clouds. Exceptions are observed for very weak HB complexes, however – volume increments have sometimes been observed for this kind of atom [6, 20, 21]. According to the topological properties it must be noticed that, for these examples, the interaction should disappear if there are small changes in the geometry.

The literature contains few studies dealing with the dependence of these properties on intermolecular distance. For complexes in the equilibrium geometry formed with dihydrogen bonds it has been observed that changes in energy, charge, and volume of the protic hydrogen tend to decrease smoothly as the H···H distance increases, following exponential relationships that depend on the donor moiety [21]. The behavior of these integrated magnitudes along reaction coordinates depends on the complex, as shown by the qualitatively different dependencies with the A···H distance for $(FH)_2$ and $(H_2O)_2$ (Fig. 16.4) [22]. Thus, for $(FH)_2$ the energy of the hydrogen atom increases regularly as the distance is reduced, up to a local maximum that is observed at a distance slightly larger than the equilibrium distance R_{eq}. Similar behavior is also observed for the integrated charge of the hydrogen atom in the H···F interaction of $(FH)_2$, but here the max-

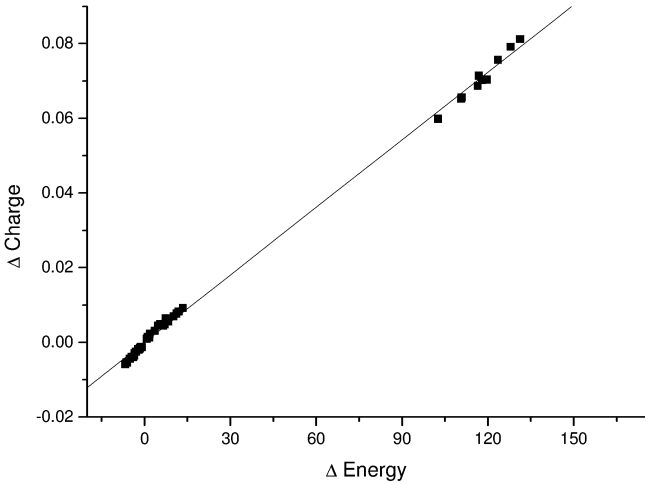


Fig. 16.3 Relationship between the change in of the net charge (e) and the change in energy (kJ mol^{-1}) for all the hydrogen atoms of tetrahydroimidazo[4,5-*d*]imidazole on dimer formation. (Values taken from Ref. [19]). The fitted linear equation is: $\Delta\text{Charge} = 0.603(4) \times 10^{-3} \times \Delta\text{Energy} - 0.05(22) \times 10^{-3}$.

imum occurs at the equilibrium distance (Fig. 16.4). These features are not observed for $(H_2O)_2$, a result that is explained by the incipient formation of a covalent bond in $(FH)_2$ that should be able to stabilize the hydrogen in $(FH)_2$, whereas for $(H_2O)_2$ this process would be hindered by stronger repulsion, leading to the more important variation observed for the integrated properties of its involved hydrogen atom (notice the different scale of Figs 16.4a and 16.4b).

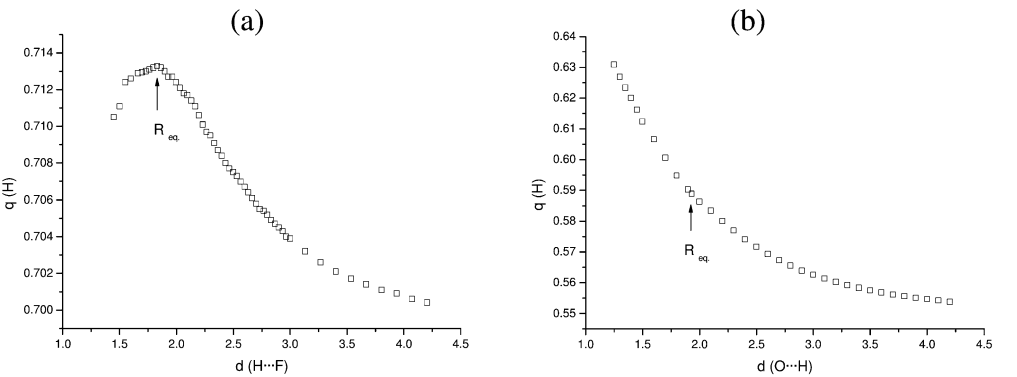


Fig. 16.4 Relationship between the integrated charge (e) on hydrogen atoms involved in H···F (a) and H···O (b) hydrogen-bonding interactions and the interaction distances (Å) for the $(FH)_2$ and $(OH_2)_2$ dimers. (Values taken from Ref. [22]).

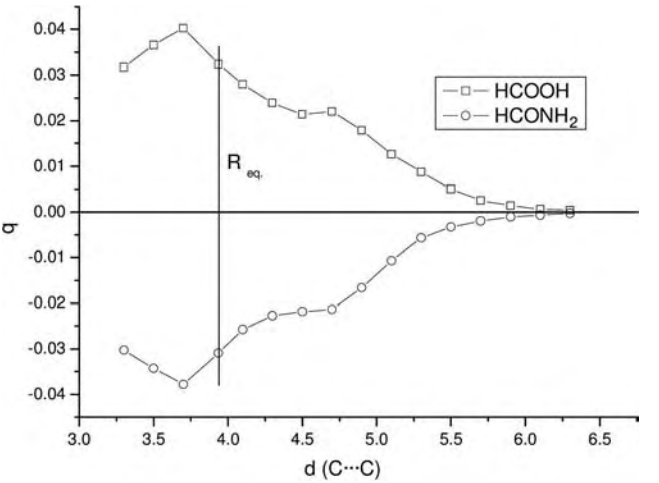


Fig. 16.5 Variation of the net molecular charge (e) with intermolecular C···C distance (Å) in the formamide/formic acid complex (values taken from Ref. [23]). R_{eq} represents the C···C distance corresponding to the equilibrium geometry of the complex.

The increase of the hydrogen net positive charge in the H···A interaction is accompanied by a decrease in the net charge of the donor molecule, indicative of the well-known electron transfer from the acceptor molecule that reproduces the expected acid–base behavior. In general, this electron transfer tends to be less important as the interatomic distance increases, being almost negligible for weaker interactions. In complexes involving a single hydrogen bond the charge transfer increases regularly as the molecules approach [22] whereas for complexes involving several hydrogen bonds more complex behavior is observed. This is true for the formamide–formic acid complex [23], which is stabilized by two HB interactions (N–H···O and O–H···O, the former being only observed for C···C distances shorter than 5.0 Å), and where both molecules behave as donor and acceptor. Despite the intricate dependence of charge transfer between molecules (Fig. 16.5), the strongest O–H···O interaction causes electron transfer from the formamide toward the formic acid in all the interaction range of distances. This transfer becomes more important as the intermolecular C···C distance shrinks to approximately 3.7 Å, which corresponds to a shorter distance than the equilibrium geometry.

16.3
Energy Properties at the Bond Critical Point (BCP)

The total electron energy density at a given point of space, $H(\mathbf{r})$, is defined as the sum of the kinetic $G(\mathbf{r})$ and the potential $V(\mathbf{r})$ contributions. The last two are

related to the local energy contribution of the Laplacian of $\rho(\mathbf{r})$ by the local form of the virial theorem (Eq. 1 in a.u.):

$$\frac{1}{4}\nabla^2\rho(\mathbf{r}) = 2G(\mathbf{r}) + V(\mathbf{r}) \tag{1}$$

Because $G(\mathbf{r}) > 0$ and $V(\mathbf{r}) < 0$, local depletion of the charge density ($\nabla^2\rho(\mathbf{r}) > 0$) is related to a preponderance of twice the kinetic energy density over the potential energy density, whereas the local concentration ($\nabla^2\rho(\mathbf{r}) < 0$) corresponds to the opposite situation. These local energy magnitudes cannot be derived from the electron density distribution alone; knowledge of the wavefunction is also necessary. Thus, the exact magnitudes of $G(\mathbf{r})$ and $V(\mathbf{r})$ (and therefore also that of $H(\mathbf{r})$) cannot be extracted from experimental analysis of $\rho(\mathbf{r})$ in crystals. They can, however, be estimated by use of Eq. (1) and Abramov's functional (Eq. 2 in a.u.), which links the topological properties of the electron distribution ($\rho(\mathbf{r})$, $\nabla\rho(\mathbf{r})$ and $\nabla^2\rho(\mathbf{r})$) to the local kinetic energy density [11, 24, 25]:

$$G(\mathbf{r}) = \left(\frac{3}{10}\right)(3\pi^2)^{2/3}\rho(\mathbf{r})^{5/3} + \left(\frac{1}{72}\right)\frac{|\nabla\rho(\mathbf{r})|}{\rho(\mathbf{r})} + \left(\frac{1}{6}\right)\nabla^2\rho(\mathbf{r}) \tag{2}$$

At BCP positions the second term of Eq. (2) vanishes, because the first derivative becomes zero, and the local energy densities are functionals of ρ_b and $\nabla^2\rho_b$. It should be noted that in the original work of Abramov, good agreement with Hartree–Fock calculations of $G(\mathbf{r})$ is obtained in the medium-range region only, i.e. for distances of approximately 0.5–2.1 Å from atomic nuclei.

Equations (1) and (2) were applied for the first time to a large experimental data set of 83 X–H···O (X = C, N, O) hydrogen-bonding interactions for which BCPs were experimentally observed between 0.5 and 1.2 Å from the hydrogen atom and between 0.5 and 1.6 Å from the oxygen atom [24]. For this data set the estimated values of G_b and V_b were exponentially dependent on the H···O distance (Fig. 16.6), and there was an exponential interdependence between them. Later, similar features were also found for H···F hydrogen-bonded systems by use of theoretical data [15].

Several theoretical articles have dealt with the validity of these local energy estimates and their associated limitations. Analysis along the bond paths of the HB complexes formed by $(FH)_2$ and $(OH_2)_2$ in their minimum energy configuration show that the Abramov functional (Eq. 2) elegantly reproduces the magnitude of the electron kinetic energy density well in the intermediate H···A region where the BCP is located but fails in regions close to the nuclei, especially around the hydrogen nucleus [26]. For a set of 37 H···F HBs, comparison of the G_b magnitudes calculated by use of Eq. (2) with those obtained from the ab-initio wave function reveals an almost perfect match [27]. In the same study, estimated values of the potential component, V_b, and the total energy density H_b, derived

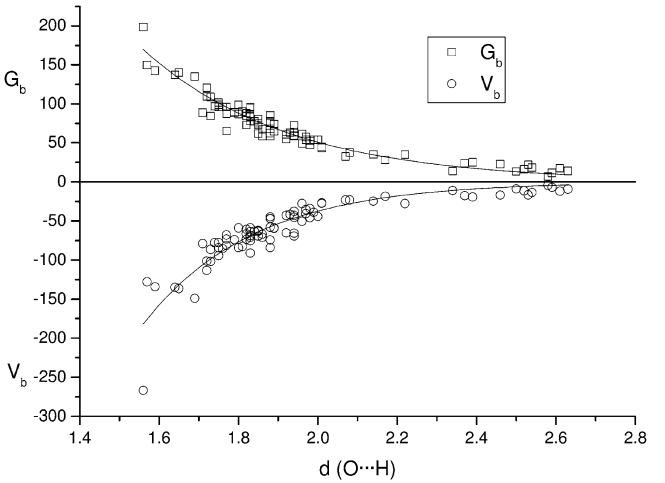


Fig. 16.6 Exponential relationships between G_b and V_b (kJ mol^{-1} a$_0^{-3}$) and the H···O distance (Å). (Values taken from Ref. [24]). The fitted curves are $G_b = 12(2) \times 10^3 \exp[-2.73(9)d(\text{O}\cdots\text{H})]$ and $V_b = -50.0(1.1) \times 10^3 \exp[-3.6d(\text{O}\cdots\text{H})]$.

from application of the virial theorem were, again, both in very good agreement with the corresponding theoretical magnitudes. Study of N–H covalent and N···H hydrogen-bonding interactions shows that the Abramov functional works well for HBs longer than 1.9 Å, overestimating the ab-initio results for closer interaction distances and providing random results for covalent N–H bonds [28]. For dihydrogen bonded complexes it has been found that Eq. (2) underestimates G_b magnitudes at long distances and overestimates them at short distances by up to 10% [21].

In addition to the exponential dependencies on H···O distance, excellent linear relationships between V and the sum of the perpendicular curvatures, $\lambda_1 + \lambda_2$, and between G and the curvature along the bond path direction, λ_3, have been found for both experimental H···O and the theoretical H···F data sets. It is particularly worthy of note that the linear regressions performed with these data sets have equivalent fitting data (Fig. 16.7). The topological and energy dependencies on H···A distance observed for the reported experimental data have enabled interpretation of the strengthening of the HB interaction in terms of those $\rho(\mathbf{r})$ properties [14, 24]. Local energy densities are dimensionally equivalent to a pressure. Hence $V(\mathbf{r})$ is interpreted as the pressure exerted on the $\rho(\mathbf{r})$ distribution to concentrate it and $G(\mathbf{r})$ is interpreted as the counterpart pressure exerted by $\rho(\mathbf{r})$ against the environment, as a reaction to the former because of the electron–electron repulsion. Thus, the greater the potential energy density the stronger the repulsion, leading to a greater kinetic energy.

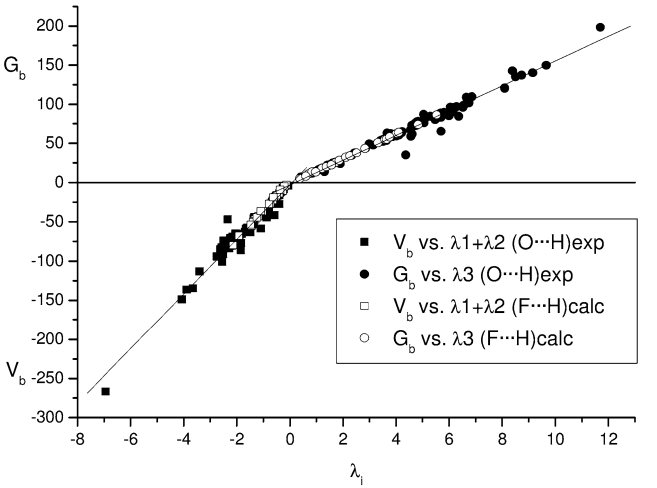


Fig. 16.7 Linear relationships between local energy densities G_b and V_b (kJ mol^{-1} a_0^{-3}) and electron density curvatures (e Å^{-5}) for experimental O···H and calculated F···H hydrogen bonds. (Values taken from Refs [14] and [15]). The linear fitting corresponds to the equations $V_b = 35.1(7)(\lambda_1 + \lambda_2)$ and $G_b = 15.3(1)\lambda_3$ for the (O···H) HB.

As a consequence of bond formation, a maximum distribution of $\rho(\mathbf{r})$ is created along the bond path; this is the indication of the established interaction. The curvatures of the electron distribution in the perpendicular plane at any point of the bond path, and in particular at the BCP, increase with increasing concentration of $\rho(\mathbf{r})$. Along the bond path $\rho(\mathbf{r})$ decreases from the nuclei towards the BCP, where it has a local minimum, and the curvature along this direction increases with charge depletion. The increase of the negative magnitude $\lambda_1 + \lambda_2$ is accompanied by an increase of the positive amplitude, λ_3, indicating that a sharper electron concentration in the plane where $\rho(\mathbf{r})$ is maximum is necessarily followed by larger $\rho(\mathbf{r})$ depletion along the direction of the bond path. Thus, at BCP, V_b and G_b are related to the charge concentration of $\rho(\mathbf{r})$ in the perpendicular plane to the bond path and to its charge depletion along the path direction, respectively. The linear relationships observed between the local energy densities and the topological curvatures reflect this situation. In this way, when the H···A distance shortens and the HB interaction becomes stronger, accumulation of charge in the internuclear region increases in magnitude with all topological and energetic properties at the BCP. Indeed, for a pure closed shell interaction, and as a consequence of Pauli's principle, the increase of ρ_b is accompanied by a more important $\rho(\mathbf{r})$ depletion ($\nabla^2\rho_b > 0$) that follows from $\lambda_3 > |\lambda_1 + \lambda_2|$ within the range of distances considered. According to the linear correlations $G_b \propto \lambda_3$ and $V_b \propto \lambda_1 + \lambda_2$, this is related to the observed ratio of pressures $|V_b|/G_b < 1$ for this type of interaction.

16.4
Topological Properties and Interaction Energy

The interaction energy, which is the stabilization occurring as a result of complex formation, provides a measure of the strength of the interaction. It is calculated as the difference between the energy of the complex and the energies of the isolated monomers, and corresponds to the negative value of the dissociation energy.

With the objective of gaining insight into the connection between the local behavior of the electron distribution in hydrogen-bonding regions and the integrated properties of complexes, the dependence of the interaction energy on the BCP properties has been explored in parallel with the distance dependencies of the latter. This approach has been undertaken to characterize the strength of hydrogen bonds in molecular crystals, thus providing a link between local $\rho(\mathbf{r})$ quantities associated with intermolecular interactions and crystal properties.

With the initially proposed linear dependencies of ρ_b and $\nabla^2\rho_b$ on hydrogen bonding distance, linear correlations of these topological properties with the interaction energy have also been reported for XCN···HCl and XCN···HF complexes (Fig. 16.8), and for carbenes as HB acceptors [8–10]. Similar results have been also recently been obtained for a wide variety of HB systems with a wide range of interaction strengths, from weakly bonded complexes in the van der Waals limit, for example CH_4···Ar, to the almost covalent interaction in the H_3O^+···H_2O complex [29]. In this last study, linear regression was applied to the ρ_b data represented against the interaction energy in the full interaction range

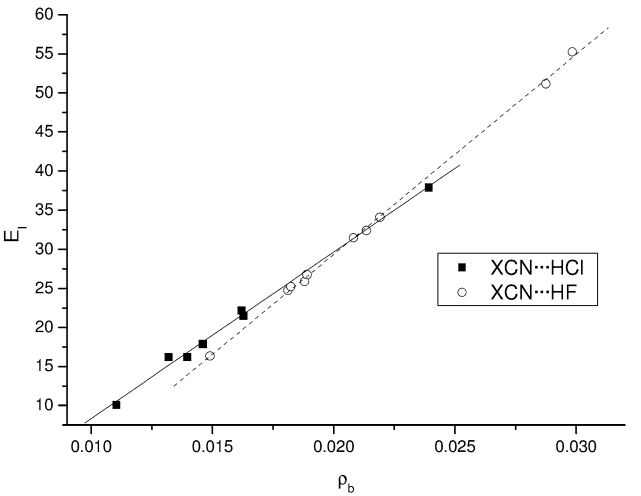


Fig. 16.8 Dependence of ρ_b (a.u.) on the interaction energy (kJ mol^{-1}) [9]. The linear regressions are: $E_i = -13(1) + 2133(66)\rho_b$ and $E_i = -21.8(6) + 2558(30)\rho_b$ for the XCN···HCl and XCN···HF complexes, respectively.

taken into account. As for $\nabla^2\rho_b$, however, although the proportionality is valid for weak HBs the regression should be regarded as an unpolished approximation for strong interactions. According to these analyses, the accumulation of electrons between the acceptor and the hydrogen, represented by ρ_b, reflects the interaction strength, characterization that has also been used for other type of interaction, for example Si–O bonds in silicates [30]. In this last example the interaction energy calculated from ρ_b confirms predictions of bond strength by empirical models.

On the basis of experimental analysis of electron density for 83 X–H···O hydrogen-bonding interactions in crystals, a different approach to the characterization of the interaction energy has been proposed in terms of the local energy densities at BCP [31]. By comparing the dependence of V_b on the HB distance observed for this set of experimental H···O interactions with that for the dissociation energies, D_e, theoretically calculated for several dimers, an approximately linear relationship between V_b and the interaction energy, $E_i = -D_e \approx 1/2V_b$ was found.

By following this phenomenological relationship, derivation of an interaction potential for H···O hydrogen bonds was undertaken [31] on the basis of the topological analysis of $\rho(\mathbf{r})$ for this experimentally characterized interaction. The interaction potential function is defined as:

$$U = -vH_b \tag{3}$$

where v is a constant in volume units (0.982 $a_0{}^3$) which was calculated by using the force constant of H···O hydrogen bonds in ice VIII ($k = 22.7$ N m^{-1}) experimentally determined from Raman spectroscopy [32]. Here H_b is expressed as the sum of the exponential dependencies fitted for G_b and V_b (Fig. 16.6). From Eq. (3) it should be noted that the potential well depth at equilibrium geometry corresponds to $(H_b)_{max}$, which represents the greatest excess of kinetic energy that ρ can afford at BCP, leading to the most efficient ρ depletion at the interatomic surface. Thus, the internuclear distance associated with $(H_b)_{max}$ (i.e. equilibrium geometry) is the best compromise between the quantity of bonding charge and its degree of depletion (reflected by ρ_b and $\nabla^2\rho_b$, respectively) furnishing the deepest stabilization of a pure closed-shell interaction [15]. According to the terms involved in Eq. (3), the interaction potential is expressed as:

$$U = 49100 \times \exp(-3.6r) - 11800 \times \exp(-2.73r) \tag{4}$$

where U is in kJ mol^{-1} and r is the H···O interatomic distance in Å. It is worthy of note that when the interaction energy function $E_i = -D_e = 1/2V_b$ has been corrected for the polarization energy $|U_0{}^{pol}|$ at their equilibrium distance, which was theoretically calculated for ice VIII [33], the new function $-D_e + |U_0{}^{pol}|$ crosses the $U = -vH_b$ potential at its minimum. As far as $-D_e(r)$ represents the interaction energy at any equilibrium distance r, the crossing between $-D_e(r) + |U_0{}^{pol}|$ and $U = -vH_b$ at U_{min} elegantly indicates the internal agreement between both descriptions (Fig. 16.9).

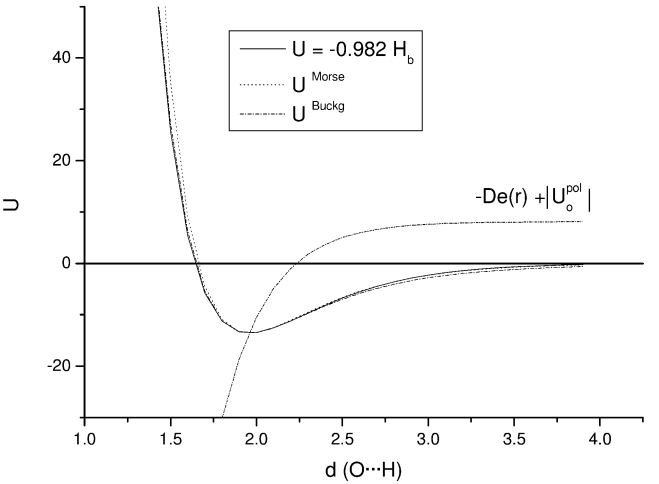


Fig. 16.9 Interaction potential $U(r) = -vH_b$ and polarization-corrected dissociation energy $D_e(r) + |U_0{}^{pol}| = -1/2V_b$ (kJ mol^{-1}) along the interatomic distance (Å), with the Morse (U^{Morse}) and Buckingham (U^{Buckg}) potentials for the hydrogen bond [31].

Both $U = -vH_b$ and $D_e = -1/2V_b$ were checked against experimentally derived thermodynamic properties of ice VIII, leading to a good estimate of the heat of sublimation at 0 K and excellent agreement with the H···O expansion distance from 0 to 273 K calculated from the linear coefficient of expansion of ice. Several energy properties, for example polarization, cohesion, and binding energies, calculated by use of those functions were also found to be in good agreement with theoretical calculations for ice VIII and ice Ih for experimental and calculated geometries. The $U = -vH_b$ potential was also successfully compared with Buckingham and Morse-type potentials, which are used in semi-empirical atom–atom potential methods and in spectroscopy, respectively, for representing hydrogen bond energy. Indeed, when they were constrained to have the same features (i.e. the same position, potential, and curvature) as those describing the potential curve $U = -vH_b$ at its minimum, the comparisons revealed almost perfect matching over the complete range of distances considered.

In a later theoretical study involving H···F hydrogen-bonding interactions [15] the proportionality factor obtained (~0.42) was similar to that derived between E_I and V_b for H···O interactions (~0.5). In this work, study of the $(FH)_2$ system at different intermolecular distances, also showed that the distance of the minimum interaction energy at equilibrium geometry also corresponds to the maximum of H_b [15]. Both theoretical findings support the $U = -vH_b$ and $D_e \approx -1/2V_b$ correspondences that were derived from experimental data. In the same theoretical study, the degree of softening term (SD), defined as H_b/ρ_b for $H_b > 0$, was also used to estimate E_i magnitudes. For the data set for neutral complexes X–H···F–Y, plots of the theoretically calculated energies E_i against the corre-

sponding values of V_b and of SD gave a narrower distribution of data for the latter when both were fitted by use of linear regression.

16.5 Electron Localization Function, $\eta(\mathbf{r})$

The electron localization function (ELF), $\eta(\mathbf{r})$, was originally defined by Becke and Edgecombe as [34]:

$$\eta(r) = 1/(1 + q^2) \tag{5}$$

where $q = (T - T_w)/T_{TF}$, and T is the kinetic energy, and T_w and T_{TF} are the von Weizsäcker and the Thomas–Fermi kinetic energy functionals, respectively. T_{TF} gives the kinetic energy of a homogeneous electron gas having the same density as the point in the electron density which is under consideration and T_w is an inhomogeneity correction. ELF takes values in the range $0 < \eta(\mathbf{r}) < 1$. (See Chapter 5 of this book for a review on the ELF).

In the regions where the electrons are localized Pauli repulsion has little effect and, therefore, $T \approx T_w$ and $\eta(\mathbf{r}) \approx 1$. If, however, the Pauli repulsion is strong and electrons are delocalized, $(T - Tw)$ is large and $\eta(\mathbf{r}) \approx 0$. The value $\eta(\mathbf{r}) = 0.5$ corresponds to the electron localization of a homogenous electron gas. ELF local maxima are known as attractors and correspond to local charge concentrations. For these attractors, the topological partition of the $\nabla\eta(\mathbf{r})$ gradient vector field yields basins that can be associated either with core electrons (core basins) or with bonds and lone pairs (valence basins). The valence basin is characterized by the synaptic order, which is the number of participating atomic valence shells. The synaptic order is given by the number of core basins sharing a boundary with the valence basin, plus the number of nuclei of hydrogen atoms it contains. Depending on their synaptic order, valence basins are classified as monosynaptic, disynaptic, trisynaptic, etc. Monosynaptic basins correspond to the lone pairs of the Lewis model, and polysynaptic basins to the shared pairs of the Lewis model. Disynaptic basins therefore correspond to two-center bonds, trisynaptic basins to three-center bonds, etc. The basins are delimited by zero-flux surfaces S $(\nabla\eta(\mathbf{r}) \cdot \mathrm{n}(\mathbf{r}) = 0,\ \forall \mathbf{r} \in S)$, where $(3, -1)$ critical points topologically analogous to BCPs are also found.

For hydrogen bonds two valence basins appear in the bonding region. One, which corresponds to the electron pair that belongs to the D–H bond, is a disynaptic basin in contact with the core basin of the donor and contains the hydrogen. The other is a monosynaptic basin that corresponds to the acceptor lone pair. Three $(3, -1)$ critical points are observed in D–H···A hydrogen bonds, two between the core and the valence basins, the other in the surface separating both valence basins. The behavior of ELF in the hydrogen-bonding region is illustrated by the ELF profiles depicted in Fig. 16.10.

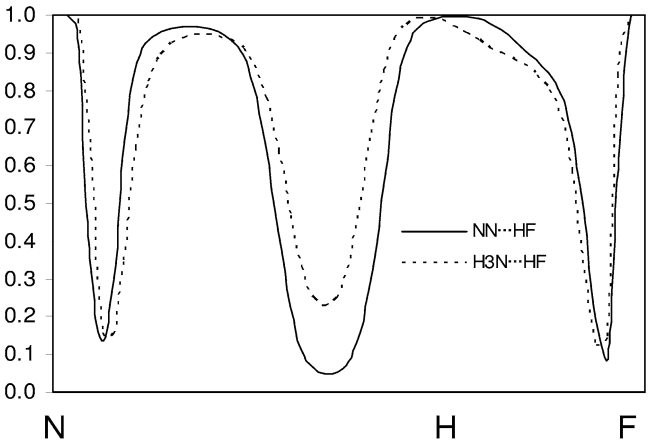


Fig. 16.10 $\eta(\mathbf{r})$ along the N···HF hydrogen bond of the N_2···HF and H_3N···HF complexes. The N···F distance has been normalized to the same value in both traces [35].

Topological analysis of the ELF has been used to characterize the strength of the HB interactions on the basis of the values of $\eta(\mathbf{r})$ at the $(3,-1)$ critical points [36]. This is done by use of the core-valence bifurcation index, which is defined as:

$$\vartheta = \eta_{vv} - \eta_{cv} \tag{6}$$

where η_{vv} is the value of $\eta(\mathbf{r})$ at the $(3,-1)$ point between both valence basins and η_{cv} is the largest of the $\eta(\mathbf{r})$ values at the two $(3,-1)$ points between the core and valence basins.

The ϑ index has been interpreted from the localization domains, which are the volumes enclosed by isosurfaces of the ELF ($\eta(\mathbf{r}) = f$, where f is a constant with $0 < f < 1$). Hence, if ϑ is negative, in the range $\eta_{vv} < f < \eta_{cv}$, there are separated localization domains for the donor and the acceptor. If ϑ is positive, however, a localization domain containing both valence attractors is observed. $\vartheta < 0$ corresponds to weak complexes, for example N_2···HF, whereas $\vartheta > 0$ is observed for stronger complexes, for example H_3N···HF (both are depicted in Fig. 16.10). In Fig. 16.10, the different sign of ϑ can be deduced from the values of ELF at the $(3,-1)$ critical points, which are the minima in the plot. For weak complexes the value of $\eta(\mathbf{r})$ at the critical point between both valence basins (η_{vv}), which is the central minimum in the plot, is below the value of $\eta(\mathbf{r})$ for the other two minima; the opposite situation is observed for the strong complexes. It has been shown that the more negative the value of ϑ, the stronger the interaction; there is an almost linear correlation between this index and the dissociation energy of the complexes [35] for a given donor (Fig. 16.11).

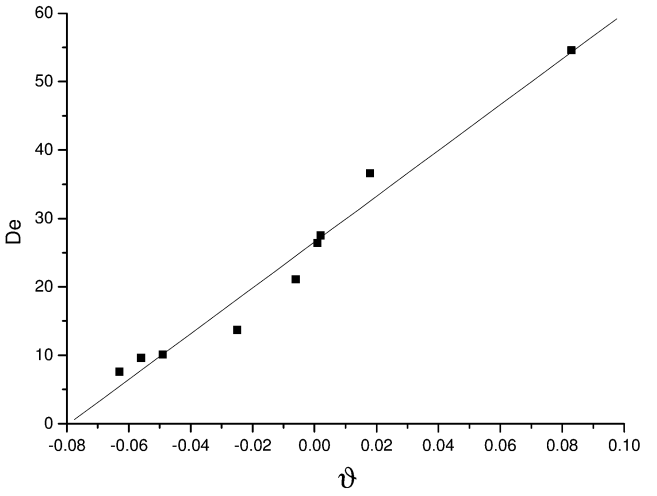


Fig. 16.11 Relationship between dissociation energy (kJ mol^{-1}) and core-valence bifurcation index (ϑ) for F–H···A complexes [36]. The fitted line corresponds to the equation $D_e = 334(22)\vartheta + 26(1)$.

The changes both in the distribution of the ELF magnitudes and in the observed domains during the proton-transfer process in strong HB complexes suggest that hydrogen bonding is accompanied by strong electron localization in the intermolecular regions. This has been shown for proton transfer between an imidazole and a carboxyl group [37]. At the middle stage of the N···H···O transfer, a single localization domain with an ELF value larger than 0.5 bridges both molecules through the hydrogen, indicating strong electron localization throughout the hydrogen-bonding region. This situation contrast with the initial and the final stages of proton transfer, in which the donor and acceptor valence shells seem to be separated by a region of depleted electron distribution, represented by ELF values of ~0.2. Three ELF attractors are, moreover, observed in the middle stage of the proton transfer, because the hydrogen is no longer included in the valence basins of either the nitrogen or the oxygen but appears as a sharp ELF peak at the hydrogen position.

16.6 Complete Interaction Range

16.6.1 Dependence of Topological and Energy Properties on the Interaction Distance

From weak van der Waals to strong covalent interactions, hydrogen atoms interact and bind with other atoms in very different forms. This can be easily observed

by making use of the topological properties of $\rho(\mathbf{r})$. For example, for H···X hydrogen-bonding interactions $\nabla^2\rho_b$ is positive whereas for X–H covalent bonds $\nabla^2\rho_b$ is negative. Thus, according to topological characterization of $\rho(\mathbf{r})$, $\nabla^2\rho_b < 0$ and $\nabla^2\rho_b > 0$ are, respectively, indicative of closed-shell and shared-shell interactions. Despite these differences, the existence of very short HB complexes has enabled almost continuous consideration of the binding properties of the pair of atoms X and H from covalent to very weak interactions.

The first attempt to find a unique dependence of ρ_b on the interaction distance, including covalent and HB complexes, corresponds to the C–H and C···H interactions. For these a semi-logarithmic relationship was proposed for 33 systems with internuclear distances between 1.0 and 2.8 Å (correlation coefficient $r^2 = 0.996$) [38]:

$$d(\text{C–H}, \text{C}\cdots\text{H}) = 0.52(2) - 0.43(3)\ \ln(\rho_b) \tag{7}$$

This equation can be also expressed as an exponential dependence of ρ_b on the distance. Double logarithm functions, which are equivalent to power functions $\rho_b = \text{A}d^{\text{B}}$ have also been successfully used to fit similar dependencies of ρ_b in covalent bonds, however [12]. Statistical analysis of the best fitting of ρ_b data for 16 types of bond, including covalent and hydrogen bonds, and using either a double logarithm or single logarithm model, has been reported [39]. According to the results, the second model, which corresponds to exponential dependence, provides slightly better statistical results for all kind of interaction.

Although a single exponential enables good fitting for the complete range of distances, statistical analysis conducted later for a large set of N–H···N complexes shows that a significantly better fit results when two separate exponential regression functions of the same type are used, one for the covalent N–H and the other for the H···N hydrogen-bonding interactions [40]. In the same way, in a theoretical study of F···H hydrogen bonded complexes [15], in which covalent F–H bonds were also included, it was proposed that data corresponding to the shared-shell and to the closed-shell regions are better fitted with independent exponentials. Then, to describe the whole interaction range with a single continuous dependence, a joint function (Eq. 8) linking both exponentials was derived. By following this method the joint function is obtained by dividing the complete range of distances into three parts, which correspond to the covalent, transition, and noncovalent interactions. After fitting of the data with single exponentials within the covalent and noncovalent regimes independently, the data included in the transition regime are fitted to a three-variable function defined as:

$$J(x) = f(x)\frac{1}{1 + e^{E(x-a)/c}} + g(x)\frac{1}{1 + e^{-E(x-b)/c}} \tag{8}$$

where a, b, and c are the fitting terms, E is a normalization factor, and $f(x)$ and $g(x)$ are the fitted exponentials within the covalent and noncovalent ranges, respectively. Further studies of covalent bonds and hydrogen-bonding interactions

involving the pairs H and X = C, N, S confirmed the quality of this method for representing the whole interaction range with a single function not only for ρ_b but for all the topological and energy properties of $\rho(\mathbf{r})$ at BCP [5, 41, 42]. Most recently, and making use of the elegantly matching Morse-type function and the potential energy function defined in terms of H_b [31], the joint function has been used to fit H_b and $\nabla^2\rho_b$ data [43].

Before the joint function was proposed, the behavior of ρ_b in the transition region between closed-shell and covalent interactions was studied for protic systems A···H···D [44]. By displacing the hydrogen atom along the reaction coordinates the behavior of ρ_b between the states A–H···D and A···H–D was characterized. The results showed that ρ_b at both critical points are related by:

$$\frac{\rho_b(1)}{\rho_b(01)} + \frac{\rho_b(2)}{\rho_b(02)} = 1 \tag{9}$$

where $\rho_b(01)$ and $\rho_b(02)$ are, respectively, the ρ_b values for the free donors A–H and D–H, and $\rho_b(1)$ and $\rho_b(2)$ are the corresponding ρ_b magnitudes in the complex. This expression, which can be derived from Eq. (7), shows that the ρ_b magnitude of interactions involving hydrogen atoms is an additive property when it is expressed as a relative quantity without dimensions. Accordingly, the relationship between the ρ_b magnitudes at both critical points in the complex is linear. Indeed, in theoretical analysis of the electron distribution in N–H···N HBs, this linear relationship was revealed between the ρ_b magnitudes corresponding to the covalent N–H bond and to the hydrogen bonding N···H interaction [28]:

$$\rho_b(\text{N–H}) = 0.3506 - 1.302\rho_b(\text{N}\cdots\text{H}) \tag{10}$$

The same study showed that the relationship between the curvatures at both sides of the hydrogen atom is much more complex than is found for ρ_b. Thus, good quality fitting could only be achieved for the ratios $\lambda_3(\text{N–H})/\lambda_3(\text{N}\cdots\text{H})$ and $\lambda_{12}(\text{N–H})/\lambda_{12}(\text{N}\cdots\text{H})$, where $[\lambda_{12} = (\lambda_1 + \lambda_2)/2]$, but for neither $\lambda_3(\text{N–H})$ nor $\lambda_{12}(\text{N–H})$ as functions of $\lambda_3(\text{N}\cdots\text{H})$ or $\lambda_{12}(\text{N}\cdots\text{H})$. Considering logarithmic dependencies, these relationships are expressed as:

$$\ln\left(\frac{\lambda_{12}(\text{N–H})}{\lambda_{12}(\text{N}\cdots\text{H})}\right) = -1.237 - 1.939\ln[-\lambda_{12}(\text{N}\cdots\text{H})] - 0.1389\ln[-\lambda_{12}(\text{N}\cdots\text{H})]^2 \tag{11}$$

and

$$\ln\left(\frac{\lambda_3(\text{N–H})}{\lambda_3(\text{N}\cdots\text{H})}\right) = 0.9033 + 1.5434\ln\{\ln[\lambda_3(\text{N}\cdots\text{H})]\} \tag{12}$$

where $r^2 = 0.998$ and $r^2 = 0.988$ for Eqs (11) and (12), respectively.

Despite the different relationships of ρ_b for closed-shell and for shared-shell interactions, it is difficult to recognize the range of distances where the transition between these two regions occurs, because of the smooth changes observed for this topological property. A similar situation is found for λ_{12}, as shown for the isolated H···F system and for a set of 79 X–H···F–Y complexes (Fig. 16.12a) [15]. Considering the parallel curvature, λ_3, the change of dependence associated

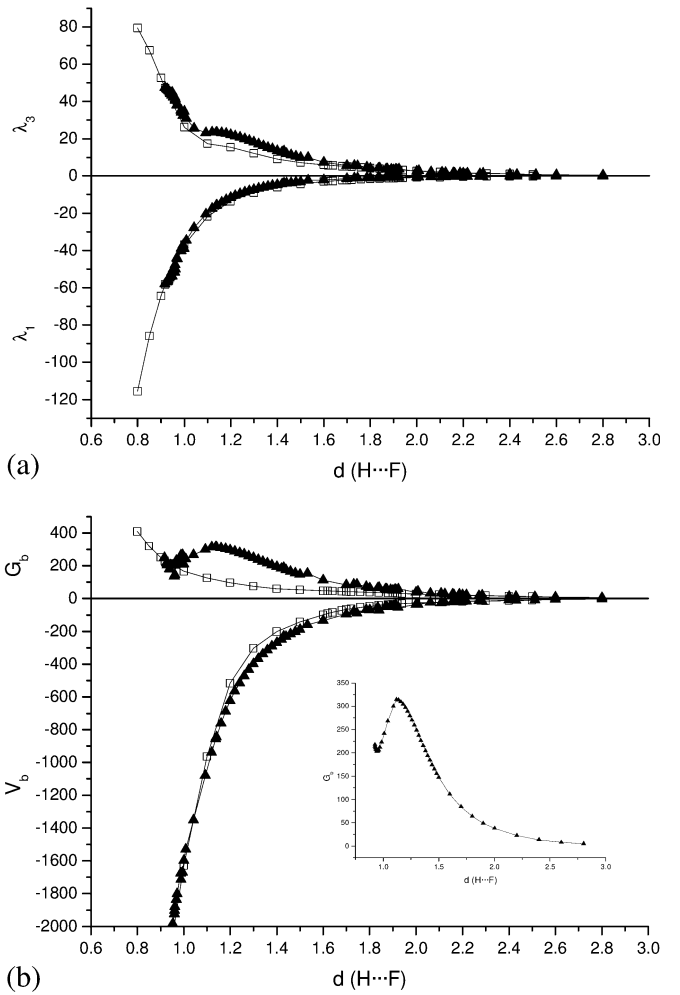


Fig. 16.12 Relationships between (a) curvatures (e Å^{-5}) and (b) local energy densities (kJ mol^{-1} a_0^{-3}) at BCP and H···F distance (Å) for the isolated H···F interaction (white symbols) and for a set of 79 X–H···F–Y complexes (black symbols). The lines were obtained by fitting joint functions [15]. The inset focuses G_b data corresponding to the [F–H···F]-system.

with the transition between both kind of interaction is revealed by a shoulder for the isolated H···F pairwise interaction and by a broad peak for the set of X–H···F–Y complexes (Fig. 16.12a). The latter corresponds to calculated values for the $[F\cdots H\cdots F]^-$ system along the reaction coordinates from $[F–H\cdots F]^-$ to $[F\cdots H–F]^-$. This proton-transfer process is energetically characterized by G_b (Fig. 16.12b), which has a local maximum at the distance $d(F\cdots H)= 1.120$ Å and is very close to that of the symmetrical proton position $[F\cdots H\cdots F]^-$ ($d(F\cdots H) = 1.138$ Å). The dependencies of the curvatures and the local energies on $d(F\cdots H)$ are similar when comparing λ_3 to G_b and λ_1 to V_b for both data sets.

Of all the topological properties at BCP, the Laplacian $\nabla^2\rho_b$ is the only one with qualitatively different behavior in covalent and noncovalent regions. Thus, in the covalent region the negative magnitude of $\nabla^2\rho_b$ increases exponentially as the bond distance decreases, because of the larger charge concentration. For a typical hydrogen-bond region $\nabla^2\rho_b$ is positive and increases as the atoms approach, because the interaction remains of the closed-shell type, and, therefore, greater charge depletion of $\rho(\mathbf{r})$ is observed.

Several studies involving reaction coordinates of the molecular approach covering positive and negative regions of $\nabla^2\rho$ have shown the generic shape of this topological property along the interatomic distance (Fig. 16.13) [15, 26]. A maximum of $\nabla^2\rho_b$ is always observed in the transition region between typical hydrogen bonds and covalent interactions, and can be fitted by making use of a joint function (Eq. 8). A similar profile, also with a maximum of $\nabla^2\rho_b$, has been ob-

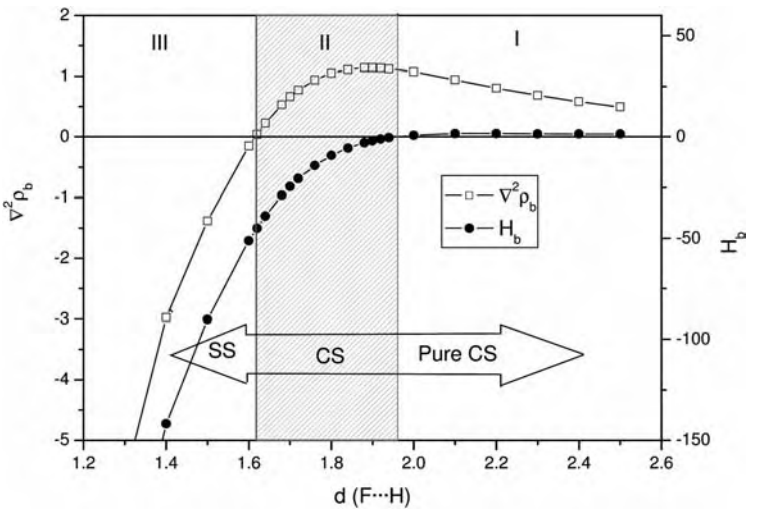


Fig. 16.13 Dependence of $\nabla^2\rho_b$ (e Å^{-5}) and H_b (kJ mol^{-1} a_0^{-3}) on the interatomic distance (Å) for isolated H···F systems [15].

served in transitions from closed-shell to shared-shell interactions involving atoms other than hydrogens, as for example LiF or CO interactions [45]. Further studies using a set of complexes at their equilibrium geometries, some of which have interaction distances in the transition region, confirmed the profile of the dependence of $\nabla^2\rho_b$ on internuclear distance [5, 15, 41, 42]. The dihydrogen bond has been also studied [46]; it has a similar $\nabla^2\rho_b$ profile, indicating that for very strong hydrogen bonds the interaction is covalent. Indeed, the negative magnitude of the Laplacian and results from energy decomposition analysis have shown that, in contrast with medium and weak hydrogen bonds, the H···H electrostatic interaction is no longer the largest attractive term for the strong hydrogen bond.

The dependence of the total electron energy density H_b on the interaction distance has a profile similar to that of $\nabla^2\rho_b$, also having a local maximum. Whereas for weak hydrogen bonds H_b is positive, because of the excess of kinetic energy G_b over potential energy V_b, for covalent bonds H_b is negative, because of the reverse situation, and becomes more negative as the atoms approach each other, leading to a greater charge concentration in the interatomic region. For intermediate examples of stronger hydrogen-bonding interactions, in which covalent features start to appear, H_b is found to be negative while $\nabla^2\rho_b$ is still positive. Positive and negative values of H_b, respectively, have been used as an alternative way of defining ionic and covalent bonds, because this quantity avoids the problems observed for $\nabla^2\rho_b$ with some covalent bonds [47]. Bonds containing very electronegative atoms (for example F_2, CO, H_2CO, and HCN) have positive values of $\nabla^2\rho_b$ whereas the corresponding values of H_b for these molecules are negative, as in other covalent bonds.

All the topological and energy properties of $\rho(\mathbf{r})$ at BCP, except $\nabla^2\rho_b$ and H_b, change smoothly with distance, varying from closed-shell to shared-shell characteristics in a continuous way and without any sudden alteration that could be associated with a transition from hydrogen bonding to covalent interactions. This transition has been studied in strongly hydrogen-bonded complexes in which proton transfer occurs [37]. For these systems, however, neither $\nabla^2\rho_b$ nor H_b have maxima in the range of distances considered, because only strong hydrogen bonds and covalent interactions are found along the proton-transfer process and the $\nabla^2\rho_b$ and H_b data calculated along the reaction coordinates were fitted to a single exponential function plus an independent term.

A coefficient which could be used to monitor closed-shell or shared-shell character and the strength of the interaction would be useful for qualitative and quantitative characterization of bonds. The ratio G_b/ρ_b was initially proposed as a classification criterion for distinguishing between different types of chemical bond [48]. Covalent interactions involve a large amount of electron density and, because the charge is locally concentrated, the kinetic energy density is expected to be comparatively small; the opposite situation is expected for closed-shell interactions, for example ionic or hydrogen bonding or van der Waals interactions. Thus, $G_b/\rho_b < 1$ and $G_b/\rho_b > 1$ have been proposed for classifying shared and

closed-shell interactions, respectively. In a theoretical study involving N–H···N hydrogen bonds with internuclear distances up to 2.3 Å, however, the G_b/ρ_b ratio never reached unity [28].

When assembling X–H and X···H interactions the characteristic features shown by both $\nabla^2\rho_b$ and H_b have been used to divide the whole range of interaction distances into three regions (Fig. 16.13) [15, 49]. Region I corresponds to weak hydrogen bonds with interaction energies $E_i < 12.0$ kcal mol^{-1} and larger interaction distances. Within this range of distances $\nabla^2\rho_b > 0$ and $H_b > 0$, and the interaction can be regarded as pure closed-shell type. The maximum of H_b occurs in this region, and the distance where $H_b = 0$ marks the border between regions I and II. Region II is situated at intermediate distances and is characterized by $\nabla^2\rho_b > 0$ and $H_b < 0$. This region is associated with a closed-shell interaction with some covalent character and contains from medium to strong hydrogen bonds with interaction energies typically in the range $12.0 < E_i < 24.0$ kcal mol^{-1}. The maximum of $\nabla^2\rho_b$ is observed in this region and is a direct consequence of the excess of V_b over G_b, starting the concentration of $\rho(\mathbf{r})$ and, therefore, being at the origin of the loss of exponential behavior of $\nabla^2\rho_b$ in region I. The transition to region III occurs at $\nabla^2\rho_b = 0$, where the shortest distances and the strongest interaction energies ($E_i > 24.0$ kcal mol^{-1}) are found. In this region, $\nabla^2\rho_b < 0$ and $H_b < 0$, both falling to deep negative magnitudes with shortening of the interaction distance. Region III corresponds to strong HBs, low-barrier hydrogen bonds (LBHB), and covalent bonds. In accordance with the topological and energy properties of $\rho(\mathbf{r})$, all these interactions are of shared-shell type.

From the local form of the virial theorem and the definition of H_b, the three regions can also be defined by the magnitudes of the ratio $|V_b|/G_b$. Thus, regions I, II, and III correspond to $|V_b|/G_b < 1$, $1 < |V_b|/G_b < 2$ and $|V_b|/G_b > 2$, respectively. This ratio enables identification and quantification of the closed-shell and shared-shell characteristics of interactions. In particular, this index has been used to classify metal–oxide bonded interactions [50] and to investigate the nature of metal–metal interactions [51].

Calculations, using the natural bond orbital (NBO) method, performed for the population of the bonding molecular orbitals for the isolated H···F system, support the partition of the whole interaction range of distances into three regions (Fig. 16.14). Indeed, according to these data, the formation of the bonding molecular orbital starts (as observed from the first converged calculation of the orbital) when the hydrogen and the fluorine atoms approach at $d(\text{F}\cdots\text{H}) \approx 2.1$ Å. This distance corresponds to the change in the behavior of the ratio $|V_b|/G_b$ (inset in Fig. 16.14), indicating a reorganization of $\rho(\mathbf{r})$ associated with a more important increase of $|V_b|$ relative to G_b, and it occurs close to that corresponding to $|V_b|/G_b = 1$, defining the border between regions I and II. For shorter distances there is rapid filling of the bonding orbital, which extends along the narrow region II. The bonding orbital is almost filled at the border of regions II and III, where $|V_b|/G_b = 2$. Further shrinkage of the bonding distance has very small

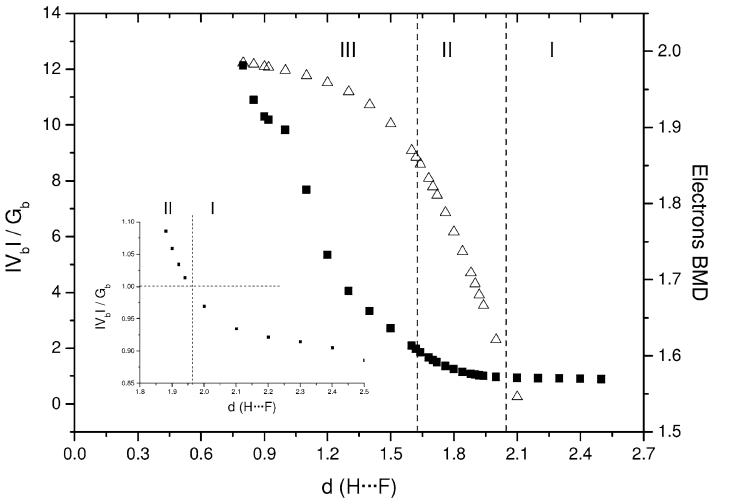


Fig. 16.14 Dependences of the index $|V_b|/G_b$ (squares) and the number of electrons filling the bonding molecular orbital (triangles) on the interatomic distance (Å) for the isolated H···F system [15].

consequences on the bonding-orbital population, which remains within the values expected for a covalent bond.

The index $|V_b|/G_b$ and the bond-degree (BD), defined as H_b/ρ_b, have been derived from analysis of the H···F interaction with the objective of characterizing pairwise atom–atom interactions [15, 51]. BD, which can be interpreted as either the total pressure per electron density unit or the total energy per electron at BCP, has been defined as the softening degree (SD) for $H_b > 0$ (Section 16.4) and as the covalence degree (CD) for $H_b < 0$. For $|V_b|/G_b < 1$ (i.e. for $H_b > 0$), the electron distribution is ideally depleted ($\nabla^2\rho_b > 0$) according to the local form of the virial theorem. As the internuclear distance shortens, the ratio $|V_b|/G_b$ increases and the interaction becomes stronger, leading to more charge in the internuclear region (ρ_b increases) and to a regular decrease of SD to zero at $|V_b|/G_b = 1$. At shorter distances, $1 < |V_b|/G_b < 2$ ($H_b < 0$) and, from Eq. (1), $\nabla^2\rho_b > 0$. Despite the positive value of $\nabla^2\rho_b$, the more important increase of $|V_b|$ in relation to G_b reduces the magnitude of $\nabla^2\rho_b$ to zero at $|V_b|/G_b = 2$. Then, for shorter interaction distances, an important increase of the negative magnitudes of both $\nabla^2\rho_b$ and H_b is observed, indicating that electrons are concentrated and the interaction has strong covalent character. As a consequence of the observed electron redistribution in the H···F bonding molecular orbital, an initial amount of covalence (CD $= H_b/\rho_b$, for $H_b < 0$) appears in H···F interactions for which $\nabla^2\rho_b > 0$ and $H_b < 0$, and the negative magnitude of CD increases regularly with shortening of the distance from $H_b = 0$ to very short geometries for covalent

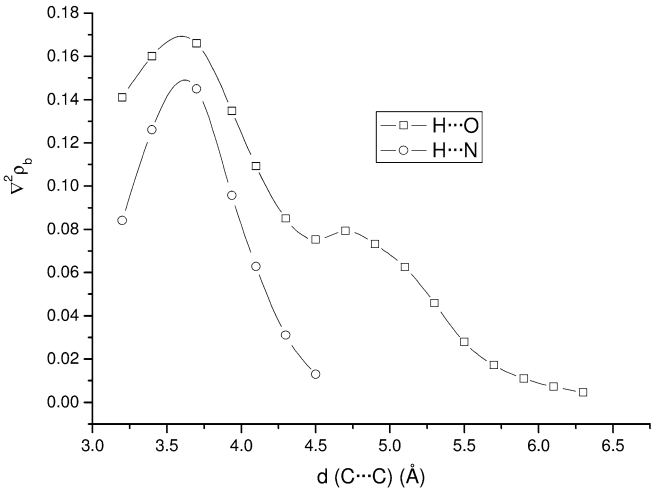


Fig. 16.15 Relationship between $\nabla^2\rho_b$ (a.u.) and the intermolecular d(C···C) distance (Å) for the HCOOH–$HCONH_2$ complex [23].

interactions ($|V_b|/G_b > 2$). Thus, BD decreases continuously and regularly from positive to negative values as the interatomic distance shortens along the entire interaction range.

16.6.2 Perturbed Systems

The dependence of the topological and energy properties of $\rho(\mathbf{r})$ at BCP throughout the complete range of interaction distances is qualitatively the same for all the A···H interactions studied. The exact quantitative form of the corresponding profiles, which can be regarded as the signature of $\rho(\mathbf{r})$, depends on the environment around the A···H interacting atoms, however. Thus, the profiles of $\nabla^2\rho_b$ and of H_b observed for different complexes that have been studied along reaction coordinates, appear displaced from one complex to another [23].

Accordingly, changes in the environment of the interaction should affect the BCP properties. Further evidence of this is given by the behavior of $\nabla^2\rho_b$ in the two hydrogen bonds present in the formamide–formic acid complex (Fig. 16.15) [23]. For large intermolecular separations a single O–H···O hydrogen bond is observed; a second N–H···O hydrogen bond is formed when the molecules are close enough. While the N–H···O interaction has the expected dependence on the reaction coordinate, a second local maximum and, consequently, a local minimum, is observed for O–H···O. The minimum is reached at the largest distance where the BCP associated with the N–H···O bond is observed, indicating that the O–H···O hydrogen bond is affected by the modification of its environment induced by formation and breaking of the N–H···O interaction.

The effect of external perturbations on the topological and energy properties at BCP has been studied by considering both addition and subtraction of one electron to and from the neutral F···H system, leading to $(F\cdots H)^+$ and $(F\cdots H)^-$ [52]. Thus, both ionic states are regarded as extreme cases of a net charge perturbation induced on the system by the environment. The most stable of the three calculated systems is the neutral one, as shown by its deepest energy, and by its largest ρ_b and $\nabla^2\rho_b$ values at the equilibrium configuration, which are indicative of greater stabilization energy and a larger and more concentrated quantity of charge in the interatomic region.

Comparing the dependence of the BCP properties on bond distance for the three systems within the range 0.8–3.0 Å enables observation of the effect of perturbation on the electron distribution of the neutral system. Thus, taking into account the electron configuration of the (F···H) neutral system $((1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi)^2(2\pi)^2(4\sigma)^0)$, in $(F\cdots H)^-$ the supplementary electron is added to the 4σ antibonding orbital. As a consequence, the extra charge expands the bonding molecular orbital toward both nuclei, increasing the depletion of the electron distribution and removing part of the charge from the bonding region, as observed in the dependencies of ρ_b and $\nabla^2\rho_b$ on bond distance, as shown in Fig. 16.16 for the latter. The presence of an additional electron also hinders formation of the bonding orbital, as seen by the highest value of the maximum of the Laplacian and its displacement toward shorter distances.

For $(F\cdots H)^+$, vertical ionization of one electron from the 2π nonbonding orbital, which is mainly localized around the fluorine, polarizes the bonding molecular orbital toward this atom, removing charge from the bonding region and, there-

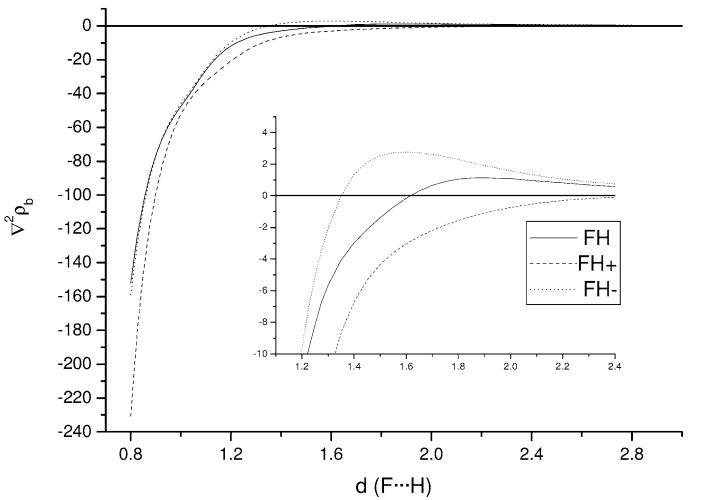


Fig. 16.16 Relationship between $\nabla^2\rho_b$ (e Å^{-5}) and the F···H distance (Å) for F···H, $(F\cdots H)^+$, and $(F\cdots H)^-$ [52].

fore, destabilizing the bond. This molecule is formed from a neutral fluorine and a proton (it has no electron charge as counterpart), so the exchange interaction does not occur within this system, leading to negative $\nabla^2\rho_b$ values and, therefore, to a net local concentration of charge over almost the full interaction range of distances. As a consequence, no local maximum of $\nabla^2\rho_b$ appears in region II (Fig. 16.16).

16.7 Concluding Remarks

This chapter contains a summary of the relationships found, so far, for the topological and energy properties of $\rho(\mathbf{r})$ in hydrogen-bonded systems, and their observed dependence on the interatomic H···A distance. The long and interesting path which awaits new research workers in this field may enable explanation of macroscopic physicochemical properties in terms of microscopic quantities derived from electron density properties. Understanding these relationships is a challenge, and a major objective in this field.

Acknowledgments

This work was supported by the Spanish Ministerio de Ciencia y Tecnología (Project No. BQU2003-01251) (IA and JE) and by the Pla de Recerca de Catalunya (Grant 2005SGR-452) (EM and IM). EE thanks Professor R. Guilard for supporting the development of a part of this work at the LIMSAG laboratory.

References

1 G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, **1997**.
2 T. Steiner, W. Saenger, *Acta Crystallogr. B* **1994**, *50*, 348–357.
3 T. Steiner, *J. Chem. Soc., Chem. Commun*. **1995**, 1331–1332.
4 M. Ramos, I. Alkorta, J. Elguero, N. S. Golubev, G. S. Denisov, H. Benedict, H. H. Limbach, *J. Phys. Chem. A* **1997**, *101*, 9791–9800.
5 M. Sánchez, P. F. Provasi, G. A. Aucar, I. Alkorta, J. Elguero, *J. Phys. Chem. B* **2005**, *109*, 18189–18194.
6 U. Koch, P. L. A. Popelier, *J. Phys. Chem*. **1995**, *99*, 9747–9754.
7 P. L. A. Popelier, *J. Phys. Chem. A* **1998**, *102*, 1873–1878.
8 R. J. Boyd, S. C. Choi, *Chem. Phys. Lett*. **1985**, *120*, 80–85.
9 R. J. Boyd, S. C. Choi, *Chem. Phys. Lett*. **1986**, *129*, 62–65.
10 I. Alkorta, J. Elguero, *J. Phys. Chem*. **1996**, *100*, 19367–19370.
11 E. Espinosa, M. Souhassou, H. Lachekar, C. Lecomte, *Acta Crystallogr. B* **1999**, *55*, 563–572.
12 O. Knop, R. J. Boyd, S. C. Choi, *J. Am. Chem. Soc*. **1988**, *110*, 7299–7301.
13 P. Roversi, M. Barzaghi, F. Merati, R. Destro, *Can. J. Chem*. **1996**, *74*, 1145–1161.
14 E. Espinosa, C. Lecomte, E. Molins, *Chem. Phys. Lett*. **1999**, *300*, 745–748.

15 E. Espinosa, I. Alkorta, J. Elguero, E. Molins, *J. Chem. Phys.* **2002**, *117*, 5529–5542.

16 P. Hobza, Z. Havlas, *Chem. Rev.* **2000**, *100*, 4253–4264.

17 X. S. Li, L. Liu, H. B. Schlegel, *J. Am. Chem. Soc.* **2002**, *124*, 9639–9647.

18 I. V. Alabugin, M. Manoharan, S. Peabody, F. Weinhold, *J. Am. Chem. Soc.* **2003**, *125*, 5973–5987.

19 I. Alkorta, O. Picazo, J. Elguero, *J. Phys. Chem. A* **2006**, *110*, 2259–2268.

20 I. Alkorta, I. Rozas, J. Elguero, *Ber Bunsen Phys Chem* **1998**, *102*, 429–435.

21 I. Alkorta, K. Zborowski, J. Elguero, M. Solimannejad, *J. Phys. Chem. A*, in press (DOI: 10.1021/jp061481x).

22 O. Gálvez, P. C. Gómez, L. F. Pacios, *J. Chem. Phys.* **2001**, *115*, 11166–11184.

23 O. Gálvez, P. C. Gómez, L. F. Pacios, *J. Chem. Phys.* **2003**, *118*, 4878–4895.

24 E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170–173.

25 Y. A. Abramov, *Acta Crystallogr. A* **1997**, *53*, 264–272.

26 O. Gálvez, P. C. Gámez, L. F. Pacios, *Chem. Phys. Lett.* **2001**, *337*, 263–268.

27 E. Espinosa, I. Alkorta, I. Rozas, J. Elguero, E. Molins, *Chem. Phys. Lett.* **2001**, *336*, 457–461.

28 O. Knop, K. N. Rankin, R. J. Boyd, *J. Phys. Chem. A* **2003**, *107*, 272–284.

29 R. Parthasarathi, V. Subramanian, N. Sathyamurthy, *J. Phys. Chem. A* **2006**, *110*, 3349–3351.

30 G. V. Gibbs, D. F. Cox, K. M. Rosso, *J. Phys. Chem. A* **2004**, *108*, 7643–7645.

31 E. Espinosa, E. Molins, *J. Chem. Phys.* **2000**, *113*, 5686–5694.

32 P. T. T. Wong, E. Whalley, *J. Chem. Phys.* **1976**, *64*, 2359–2366.

33 C. Gatti, B. Silvi, F. Colonna, *Chem. Phys. Lett.* **1995**, *247*, 135–141.

34 A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397–5403.

35 M. E. Alikhani, F. Fuster, B. Silvi, *Struct. Chem.* **2005**, *16*, 203–210.

36 F. Fuster, B. Silvi, *Theor. Chem. Acc.* **2000**, *104*, 13–21.

37 L. F. Pacios, O. Gálvez, P. C. Gómez, *J. Chem. Phys.* **2005**, *122*, 214307.

38 I. Alkorta, I. Rozas, J. Elguero, *Struct. Chem.* **1998**, *9*, 243–247.

39 I. Alkorta, L. Barrios, I. Rozas, J. Elguero, *J. Mol. Struct. (Theochem)* **2000**, *496*, 131–137.

40 O. Knop, K. N. Rankin, R. J. Boyd, *J. Phys. Chem. A* **2001**, *105*, 6552–6566.

41 O. Picazo, I. Alkorta, J. Elguero, *J. Org. Chem.* **2003**, *68*, 7485–7489.

42 I. Alkorta, O. Picazo, J. Elguero, *Tetrahedron Asymmetry* **2004**, *15*, 1391–1399.

43 P. M. Dominiak, A. Makal, P. R. Mallinson, K. Trzcinska, J. Eilmes, E. Grech, M. Chruszcz, W. Minor, K. Wozniak, *Chem. Eur. J.* **2006**, *12*, 1941–1949.

44 I. Alkorta, I. Rozas, J. Elguero, *J. Mol. Struct. (Theochem)* **1998**, *452*, 227–232.

45 J. Hernandez-Trujillo, R. F. W. Bader, *J. Phys. Chem. A* **2000**, *104*, 1779–1794.

46 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2005**, *109*, 4331–4341.

47 D. Cremer, E. Kraka, *Croat. Chem. Acta* **1984**, *57*, 1259.

48 R. F. W. Bader, H. Essen, *J. Chem. Phys.* **1984**, *80*, 1943–1960.

49 I. Rozas, I. Alkorta, J. Elguero, *J. Am. Chem. Soc.* **2000**, *122*, 11154–11161.

50 G. V. Gibbs, D. F. Cox, T. D. Crawford, K. M. Rosso, N. L. Ross, R. T. Downs, *J. Chem. Phys.* **2006**, *124*, 847041–847048.

51 G. Gervasio, R. Bianchi, D. Marabello, *Chem. Phys. Lett.* **2004**, *387*, 481–484.

52 E. Espinosa, I. Alkorta, I. Mata, E. Molins, *J. Phys. Chem. A* **2005**, *109*, 6532–6539.

17
Relationships between QTAIM and the Decomposition of the Interaction Energy – Comparison of Different Kinds of Hydrogen Bond

Sławomir J. Grabowski

17.1
Introduction

Intramolecular and intermolecular hydrogen-bonding interactions are the subject of intense interest because of their role in many physical, chemical, and biological processes, for example they affect the arrangement of molecules in crystals, the behavior of liquids and gases, and determine properties of materials such acidity, basicity, and susceptibility to electrophilic or nucleophilic substitution, etc. All of these effects of hydrogen bonding are rooted in the underlying electron density distribution, because ''matter is a distribution of charge in real space, of pointlike nuclei embedded in the diffuse density of electronic charge'' [1]. The quantum theory of atoms in molecules (QTAIM) is a powerful tool that can be used for analysis of inter-atomic interactions, as is apparent from this book.

Analysis of hydrogen bonds as a specific class of weak interactions is often performed in physicochemical studies [2–5]. QTAIM [6] uses novel descriptors, for example the properties of the electron density at the bond critical point (BCP), which enable one to gain deeper insight into the nature of the chemical bond as reviewed in this book (the properties of BCP and their interpretation are reviewed in Chapter 1).

In a typical hydrogen-bonding interaction, generally symbolized X–H···Y, where X–H designates the proton donating bond and Y is an acceptor centre, the bond critical points of both the X–H and of the H···Y bonding interactions provide crucial information characterizing the bonding. The bond properties determined at the X–H and H···Y BCPs include:

- the electron densities at the BCPs (ρ_{XH}, $\rho_{H\cdots Y}$);
- their Laplacians ($\nabla^2\rho_{XH}$, $\nabla^2\rho_{H\cdots Y}$); and
- the energy properties of both bond critical points, usually designated H_b, G_b, and V_b (Chapter 1).

In other words, QTAIM puts at ones disposal an entire set of topological properties for characterizing both bonding interactions of all types and traditional

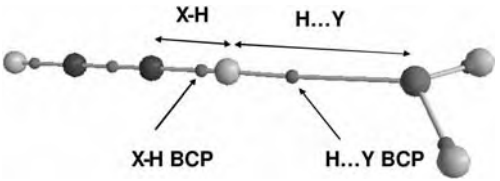


Fig. 17.1 The hydrogen bond in the acetylene–water complex.

geometrical properties. Figure 17.1 shows an example of a hydrogen-bonded system, the acetylene–water complex, with a C–H proton-donating bond and the oxygen atom of water as an acceptor, in which the topological characteristics of these interactions may be considered in addition to the X–H (C–H) and H···Y (H···O) bond lengths.

One of the early definitions of hydrogen bonding is that of Pauling [7] who stated, ''under certain conditions an atom of hydrogen is attracted by rather strong forces to two atoms, instead of only one, so that it may be considered to be acting as a bond between them. This is called the hydrogen bond.'' QTAIM enables characterization of a wide diversity of hydrogen bonding interactions (X–H···Y) by use of the above-mentioned topological descriptors, irrespective of the nature of X and Y. ''A hydrogen bond, which includes the van der Waals complexes, is defined to be one in which a hydrogen atom is bound to the acid fragment by a shared interaction, $\rho(r_c)$ large and $\nabla^2\rho(r_c) < 0$, and to the base by a closed-shell interaction, $\rho(r_c)$ small and $\nabla^2\rho(r_c) > 0$'' [8].

The wide diversity of hydrogen bonds is the subject of different classification schemes. In several monographs classification is achieved on the basis of the hydrogen-bond energies – weak hydrogen bonds (1–4 kcal mol^{-1}); medium (4–15 kcal mol^{-1}), and strong (15–40 kcal mol^{-1}) [3, 4]. It is worth mentioning that the hydrogen bond energy (E_{HB}) is often identified with the binding energy [9], which is usually computed as the difference between the total energy of the complex (E_{AB}) and the energies of the isolated monomers (E_A and E_B). E_{AB} assumes negative values for stable complexes.

$$E_{HB} = E_{AB} - (E_A + E_B) \tag{1}$$

For simplicity, the absolute values of binding energies are denoted $|E_{HB}|$ in the text and E_{HB} values are given in figures. In this chapter the *hydrogen bond energy* is the energy of the H···Y interaction *within* the X–H···Y system and the *binding energy* refers to the interacting system as a whole, as defined in Eq. (1). The terminal parts of molecules sometimes contribute substantially to binding energies. In this chapter, both terms are equivalent because:

1. interactions between terminal moieties are negligible in most systems (with a few exceptions which are described in detail); and
2. all reported hydrogen bond energies are calculated in accordance with Eq. (1).

Rozas et al. [10] have proposed a classification of hydrogen-bonding interactions based on energy and topological data. For weak hydrogen bonds with hydrogen bond energy $|E_{HB}| < 12$ kcal mol^{-1}, $\nabla^2\rho_{H\cdots Y} > 0$ and $H_b > 0$. For hydrogen bonds of medium strength 12 kcal mol^{-1} $< |E_{HB}| <$ 24 kcal mol^{-1}, $\nabla^2\rho_{H\cdots Y} > 0$ and $H_b < 0$. For strong hydrogen bonds $|E_{HB}| > 24$ kcal mol^{-1}, $\nabla^2\rho_{H\cdots Y} < 0$ and $H_b < 0$ (where H_b is the electron energy density at H···Y BCP). This classification shows that weak hydrogen bonds eventually merge with (weaker) van der Waals interactions whereas strong hydrogen bonds merge, at the other end of the continuum, with covalent and polar bonds. $\nabla^2\rho_{AB} < 0$ is indicative of covalent A–B bonds and $\nabla^2\rho_{H\cdots Y} < 0$ is also characteristic of very strong hydrogen bonds; this means that such hydrogen bonds have the characteristics of covalent bonds. (The covalence of strong hydrogen bonds is discussed elsewhere [11–13].) A negative value of H_b is often regarded as sufficient reason to classify an interaction as covalent [14]. It is, in contrast, difficult to classify an interaction as hydrogen bond if it is very weak, conditions under which hydrogen-bonding criteria are not universally accepted [15]. These very weak interactions may be regarded as van der Waals attractive interactions. One can state there is a continuous transition from covalent bonds to hydrogen bonds and, finally, to van der Waals interactions [16, 17]. For this reason, Desiraju has suggested that hydrogen bonding is an ''interaction without borders'' [18].

Figure 17.2 shows the relationship between H···Y distance and hydrogen bond energy; black circles correspond to $\nabla^2\rho_b < 0$, grey circles to $\nabla^2\rho_b > 0$ and $H_b < 0$ and white circles to $\nabla^2\rho_b > 0$ and $H_b > 0$. It is apparent that energies and H···Y distances correspond approximately to the topological properties. Hydrogen bonds are strong if the H···Y distance is less than ca. 1.2 Å and weak if the distance is greater than ca. 1.8 Å. Figure 17.2 was obtained on the basis of the different types of hydrogen bond discussed below.

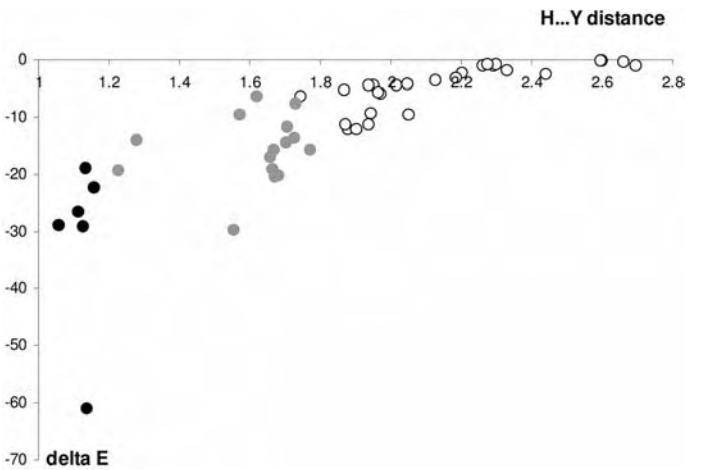


Fig. 17.2 The relationship between H···Y distance (Å) and the binding energy (kcal mol^{-1}).

Parthasarathi et al. [19] recently proposed yet another classification based on the relationship between H···Y distance and the electron density at the corresponding BCP, covering the regions from covalence to van der Waals interactions. The topological properties at the BCP can enable better characterization of a particular hydrogen bonding interaction than the binding energy, because the properties determined at the BCP are specific to the particular H···Y interaction whereas the binding energy, as already mentioned, more often than not includes, in addition to the contribution of the hydrogen bond of interest, contributions from interactions between other parts of molecules, which can sometimes be significant [20]. The BCP properties single out the characteristic of the H···Y interaction of interest from the rest of the system. Having said that, however, one must also recognize that the bond path and BCP attributed to H···Y are affected by the electron density of the whole system. An example is represented by the T-shaped configuration of the LiC≡CLi···HF complex, in which an F–H···π hydrogen bond path is present. The binding energy calculated for that system at the MP2/6-311++G(d,p) level of theory [21] amounts to 15.4 kcal mol^{-1} whereas such energy for the T-shaped HCCH···HF complex is 3.1 kcal mol^{-1}. The principal contribution to the energy for the former system seems to be Li···F electrostatic interactions.

17.2 Diversity of Hydrogen-bonding Interactions

Pauling argued that hydrogen bonds may only be formed by electronegative atoms (the X and Y atoms of X–H···Y) and stressed that the hydrogen bond is electrostatic in nature [7]. Despite Pauling's statements, early studies indicated that even atoms of very low electronegativity, for example carbon, can act as hydrogen donors in a hydrogen bond and that C–H···Y hydrogen bonds are possible [22]. Such C–H···Y hydrogen bonds were first found in crystal structures [23]; later their existence was proved by use of refined statistical methods [24]. It has been pointed out that carbon atoms may also act as proton acceptors and that X–H···C hydrogen bonds [25] and even C–H···C [26–28], are possible.

Interestingly, π-electrons can also act as proton acceptors, i.e. as Lewis bases. C–H···π hydrogen bonds are often responsible for the arrangement of molecules in crystals [29]. These unconventional hydrogen bonds are usually very weak.

The dihydrogen bond (DHB) is a special type of hydrogen-bonding interactions. The DHB may be denoted X–H$^{+\delta}$···$^{-\delta}$H–Y, because of the presence of a proton-donating X–H bond, as for a typical hydrogen bond, but, unlike a typical hydrogen bond, the proton acceptor is a second hydrogen atom bearing excess negative charge [30]. Such interactions were extensively studied in the last decade and were found to have characteristics very similar to those of typical hydrogen bonds [31]. Dihydrogen bonding is extremely important in (bio)chemistry because, for example, it occurs in the preliminary stages of the release of hydrogen gas in several biochemical processes [32].

An important characteristic of DHBs is that they are relatively strong compared with typical hydrogen bonds – the binding energy of a dihydrogen bonded complex very often exceeds 10 kcal mol^{-1} [33, 34]. For example, the binding energy for an FH···HLi model system is 11.9 kcal mol^{-1} at the QCISD(T)/6-311++G(d,p) level of approximation [33]. Very strong DHBs [13, 35], for example those in the H_2OH^+···HBeH complex and other, related, systems have binding energies in the range 20–30 kcal mol^{-1}. At the other end of the scale, very weak DHBs are also known which border van der Waals interactions [36]. In other words, DHBs cover a broad spectrum of interactions from very weak (van der Waals-like) to very strong (covalent-like) [17] in the same way as conventional hydrogen bonds.

It is crucial to know the nature of a DHB interaction, because H···H contacts are very common, for example in crystal structures of organic compounds. Such contacts may be stabilized by DHBs, by van der Waals interactions, and by hydrogen–hydrogen interactions, which have recently been detected and characterized [1].

Figure 17.3 shows the relationship between proton···acceptor distance and $-V_b/G_b$ ratio. This relationship for H···F interactions – covalent-shared and

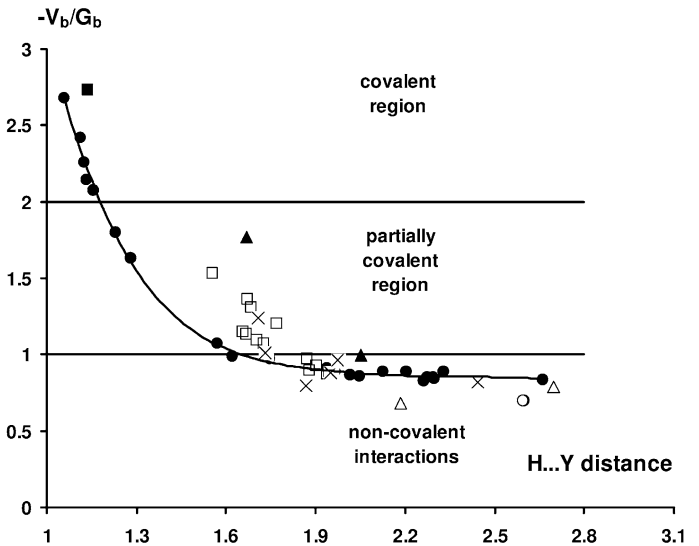


Fig. 17.3 Relationship between proton–acceptor distance and $-V_b/G_b$ ratio. Filled circles denote dihydrogen bonds. Open circles denote species with σ-electrons as proton acceptors (two species are covered in the figure because the H···Y distance and $-V_b/G_b$ ratio are close to each other). The filled square denotes $[FHF]^-$. Open squares denote so-called "resonance-assisted hydrogen bonds", for example OH···O, NH···O, and OH···N. Filled triangles denote π···H^+···π interactions. Open triangles denote complexes with π-electrons as a proton acceptor (e.g. the T-shaped acetylene dimer and the T-shaped FH···C_2H_2 complex). Crosses denote other hydrogen bonded systems; among these is the trans-linear water dimer.

non-covalent-closed-shell interactions – has been investigated [37]. The dependence presented here (Fig. 17.3) is mainly that for DHB systems (filled circles), but other types of hydrogen bond are also included. The "charge-assisted hydrogen bond" (filled square) is represented by $[FHF]^-$ and there are also so-called "resonance-assisted hydrogen bonds" [38] and, finally, typical hydrogen bonds such as that in the trans-linear water dimer (see caption of Fig. 17.3).

From Fig. 17.3 it is apparent that $-V_b/G_b$ increases if the H···Y distance (where Y denotes the proton acceptor) decreases. The (nonlinear) correlation is particularly strong for dihydrogen-bonded systems. Changes of $-V_b/G_b$ have the same trends as for other hydrogen-bonded interactions.

Because the electronic potential and kinetic energy density at the BCP (V_b and G_b, respectively) are negative and positive, respectively, everywhere the decrease in the H···Y distance results in an increase of $-V_C/G_C$. This last ratio may be treated as a measure of the covalence of a chemical bonding interaction. When the ratio is greater than 2, the corresponding Laplacian at the H···Y BCP is negative ($1/4\nabla^2\rho_b = 2G_b + V_b$), and when the ratio is between 1 and 2 the Laplacian is positive and H_b ($H_b = V_b + G_b$) is negative. Finally, when this ratio is less than 1, both values, i.e. the Laplacian and H_b, are positive. These three previously mentioned regions correspond, respectively, to covalent interactions, partially covalent interactions, and noncovalent attractive interactions, for example moderate and weak hydrogen bonds and van der Waals interactions.

Figure 17.4 shows the relationship between H···H distance and the Laplacian of the electron density at the corresponding BCP. The region of negative values of the Laplacian corresponding to very strong covalent interactions is clearly appar-

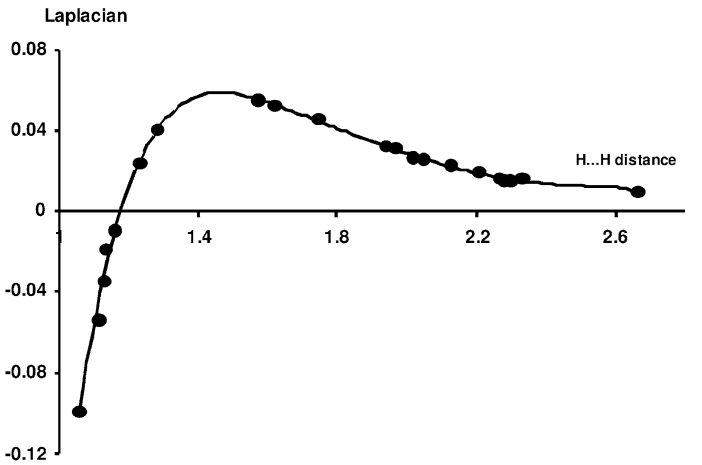


Fig. 17.4 The relationship between the H···H distance (in Å) and the Laplacian of the electron density at the bond critical point (in au) for dihydrogen bonded systems.

ent, as also is the region of positive Laplacian for the interactions of closed-shell systems; this is in agreement with Fig. 17.3.

17.3 The Decomposition of the Interaction Energy

It is claimed in different definitions of hydrogen bonding that it is an electrostatic interaction [2–4]. It is also stated, occasionally, that hydrogen bonding has a covalent nature especially when it is strong or very strong. According to the electrostatic–covalent hydrogen bond (ECHB) model [39, 40] weak and moderate hydrogen bonds are mostly electrostatic. When the proton···acceptor distance decreases, the strength of the hydrogen bond increases and so does its covalence and the electrostatic nature becomes less important.

Desiraju [18] claims that hydrogen bonding is a conglomerate of electrostatic (acid/base), polarization (hard/soft), van der Waals (dispersion/repulsion), and covalent (charge-transfer) interactions. These energy components can be obtained from an interaction energy-decomposition scheme. The approach proposed by Morokuma [41] can be used to decompose the binding energy. The Hartree–Fock energy is expressed as:

$$\Delta E_{\mathrm{SCF}} = \mathrm{ES} + \mathrm{PL} + \mathrm{EX} + \mathrm{CT} + \mathrm{MIX} \tag{2}$$

where ES is the electrostatic interaction energy, PL the polarization interaction energy, defined as the energy of the distortion of the monomers' charge distributions, and CT the energy of charge transfer between the monomers (both CT and PL interaction energy terms are referred to changes of the electronic charge distribution as a result of complexation). These terms are most often attractive for stable complexes whereas EX, the exchange energy resulting from antisymmetrization of the wave function, is usually not. MIX is the energy difference between the SCF interaction energy and these four components. Although the correlation energy (CORR) is not included in the SCF binding energy (ΔE_{SCF}), it can be calculated as the difference between the energy when correlation is taken into account (ΔE) and the SCF energy.

$$\Delta \mathrm{E} = \Delta \mathrm{E}_{\mathrm{SCF}} + \mathrm{CORR} \tag{3}$$

The dispersion energy, often attributed to the van der Waals interaction, is the most important attractive energy component within the electron correlation energy. The dispersion energy originates from mutual polarization of the electron charge distribution of interacting monomers. In other words, it is the interaction of instantaneous multipoles [42].

Results from the decomposition scheme based on the variation–perturbation approach [43, 44] are presented here. The components of the interaction energy are obtained within the MP2 method for this decomposition scheme (designa-

tions different from those of the Morokuma scheme are used here to distinguish both approaches):

$$\Delta E_{MP2} = E_{EL}{}^{(1)} + E_{EX}{}^{(1)} + E_{DEL}{}^{HF} + E_{CORR}{}^{(2)} \quad (4)$$

where $E_{EL}{}^{(1)}$ is the first-order electrostatic term describing the Coulombic interaction of static charge distributions of interacting molecules, $E_{EX}{}^{(1)}$ is the first-order exchange component resulting from the Pauli exclusion principle, and $E_{DEL}{}^{HF}$ and E_{CORR} correspond to higher-order delocalization and correlation terms. The delocalization term represents the effect of mutual deformation of the electron density as a result of complexation [45]. In other words, changes of the electron distribution within connected monomers affect both interacting species and electron transfer between them. Hence this term approximately contains CT and PL interaction energies of the Morokuma and Kitaura approach.

The delocalization interaction energy term can be related to the delocalization index (DI) [46], because the latter is a quantitative measure of the sharing of electrons between two interacting species (because it measures the number of electrons pairs delocalized between two atoms in the absence of significant charge transfer between them). The sharing of electrons between the hydrogen atom and the proton acceptor Y is usually larger the stronger the hydrogen bond [47].

According to the variation–perturbation approach (Eq. 4) the starting wave functions of the subsystems are obtained by using a dimer-centered basis set (DCBS) [48], significantly reducing the basis set superposition error (BSSE) for the total interaction energy and its components. It has been shown that this approach enables removal of the BSSE contributions to the Heitler–London first-order interaction energy term, E_{HL} [49]:

$$E_{HL} = E_{EL}{}^{(1)} + E_{EX}{}^{(1)} \quad (5)$$

It is apparent that decomposition schemes can be useful for gaining insight into the nature of the interactions of interest and, perhaps, for obtaining a definition for the term "covalence".

17.4
Relationships between the Topological and Energy Properties of Hydrogen Bonds

Different interactions are analyzed in this section in accordance with the variation–perturbation approach based on calculations at the MP2/6-311++G(d,p) level of approximation. Full-geometry optimizations were performed for all complexes considered and in this step no BSSE correction was applied to the Born–Oppenheimer (BO) energy surface. In recent studies, geometry optimizations have been performed on the BO surface corrected for BSSE. For weakly interacting monomers, some reports indicate that the differences between the corrected and uncorrected BO surfaces are negligible [50] whereas

other studies indicate that such differences can be significant [51]. For all results presented here BO energy surface was not corrected for BSSE, however, for such optimized geometry the BSSE was reduced significantly by the applied decomposition scheme described in the previous section.

Decomposition of the interaction energy yields insight on the character of the interaction. For example, for the trans-linear dimer of water the electrostatic component is -8.8 kcal mol^{-1}, the exchange interaction energy term is 6.8 kcal mol^{-1}, and the remaining delocalization and correlation terms are -2.2 and -0.3 kcal mol^{-1}, respectively. For the $H_2OH^+\cdots HBeH$ complex already discussed, decomposition of interaction energy is as follows. The electrostatic, exchange, delocalization, and correlation terms are -12.0, 19.9, -24.4, and -4.2 kcal mol^{-1}, respectively. It is apparent that for the water dimer the electrostatic interaction energy is the most important attractive term. For the much stronger dihydrogen bond in the second example the delocalization is twice as large (if one considers the modulus) than the electrostatic term. In other words one may expect for strong short hydrogen bonds that the most important attractive term is delocalization whereas for typical moderate and weak hydrogen bonds the electrostatic contribution dominates. The greater importance of delocalization is, hence, attributed to covalence. For very strong DHBs, as in the $H_2OH^+\cdots HBeH$ complex and related species, a strong correlation is found between both the exchange energy and the delocalization on one hand and the H···H distance on the other. The other interaction energy terms do not correlate with this ''intermolecular'' distance (H···H distance). Figure 17.5 shows these relations for very strong DHBs. For very strong hydrogen bonds, which are covalent in nature, the delocal-

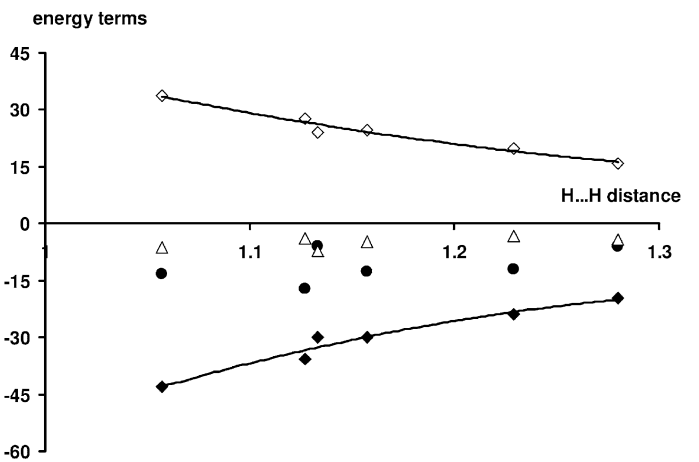


Fig. 17.5 Correlations between H···H distance (Å) and the interaction energy components (kcal mol^{-1}). Empty squares show the exchange energy, full squares the delocalization energy, full circles the electrostatic energy, and empty triangles the correlation energy.

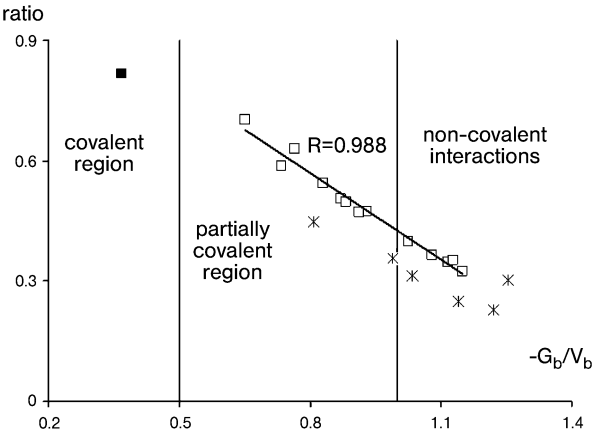


Fig. 17.6 Relationship between the $-G_b/V_b$ ratio and the ratio of the interaction energy components (delocalization to electrostatic). The meanings of the symbol are as for Fig. 17.3.

ization and exchange energies are responsible for the stability of the systems containing them. One can see from Fig. 17.5 that the importance of electrostatic interaction increases with decreasing H···H distance but not to the same extent as the delocalization energy. In other words, the ratio of the delocalization and electrostatic energy terms should correspond to the strength of hydrogen bonding. Such a correlation was found for the formamide dimer and its fluorine derivatives [52]. It has also been found [53] that when this ratio is ∼0.45 the hydrogen bonding interaction is covalent or partly covalent, as is apparent from the negative sign of the Laplacian of the electron density at the H···Y BCP or at least the negative value of H_b. This ratio, similar to the $-G_b/V_b$ ratio, is thus a measure of the strength of the hydrogen bond and of the covalence of the interaction. The relationship between both ratios is depicted in Fig. 17.6 for a homogeneous sample of bonding interactions (empty squares correspond to related systems with O–H···O and N–H···O hydrogen bonds assisted by π-electron delocalization, the full square denotes the $[FHF]^-$ system, and crosses correspond to other complexes, for example the water dimer). The linear correlation coefficient for the first class of systems is 0.988. Figure 17.6 also shows the regions of covalence, partial covalence, and of the weaker interactions for which both the Laplacian and H_b values are positive.

Figure 17.6 also shows that there is a correlation between topological data derived from the QTAIM and the energy data obtained from decomposition of the interaction energy. Hence the question arises of whether the topological data correlate with the energy components. The electronic potential energy density at the BCP correlates with the H···Y distance for hydrogen bonded systems, because it was found for O–H···O interactions that $E_{HB} \approx \frac{1}{2} V_b$ [54]. It follows that the high values of hydrogen bond energy and of the modulus of V_b are indicative of cova-

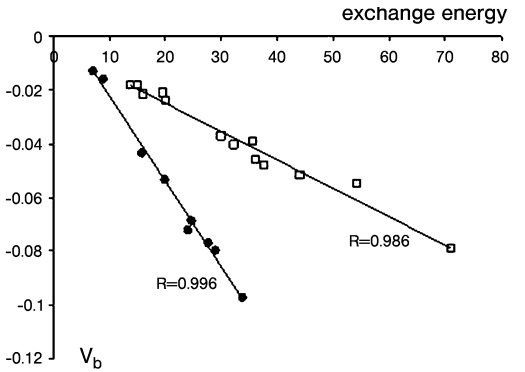


Fig. 17.7 Relationship between the exchange energy (kcal mol^{-1}) and the potential energy density at the H···Y BCP (V_b in au). Circles denote CAHB-DHBs and squares denote OH···O and NH···O.

lence because if $|V_b| > 2G_b$ then $\nabla^2\rho_b < 0$. A negative value of the Laplacian indicates a concentration of electronic charge in the inter-atomic region.

Figure 17.7 shows the correlation between the exchange energy component and V_b. Two groups are considered:

1. OH···O, N–H···O, and O–H···N (formamide and its derivatives, simple carboxylic acids); and
2. charge-assisted DHBs systems (H_2OH^+···HBeH complex and related species).

A strong correlation within each of these groups can be seen in the figure. This is connected with the strong exponential correlation between the exchange energy and the proton···acceptor distance [55]. The exchange energy, similar to V_b, correlates well with the binding energy and with the H···Y distance, the latter often reflects the strength of hydrogen bonding which (at the limit) is sometimes referred to as the "short strong hydrogen bond" (SSHB) [56].

Figure 17.8 shows the relationship between the delocalization energy and V_b, where again two groups of complexes are considered. The figure shows there is a strong correlation between the delocalization energy and V_b for both groups, as is revealed by the high regression coefficients.

Figure 17.9 shows the relationship between the electrostatic interaction energy term and V_b. It is very interesting that for the sample of systems in which there is strong π-electron delocalization (NH···O and OH···O hydrogen bonds in carboxylic acids and formamides) there is a linear relationship between both the topological and energy data. The situation is dramatically different for charge-assisted dihydrogen bonds for which interactions are very strong and H···H distances are short. In such circumstances there is no correlation. The delocalization and exchange terms correlate with V_b for very strong interactions whereas the electrostatic term does not. This indicates that the electrostatic interaction, as op-

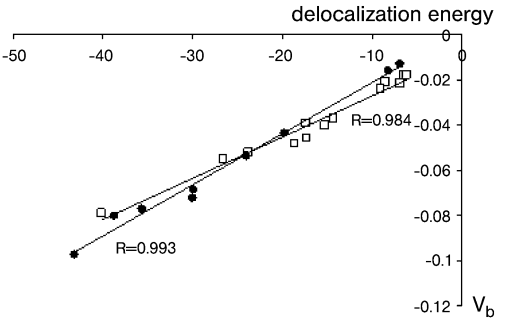


Fig. 17.8 Relationship between the delocalization energy component (kcal mol^{-1}) and the potential energy density at the H···Y BCP (V_b in au). Circles denote CAHB-DHBs and squares denote RAHBs.

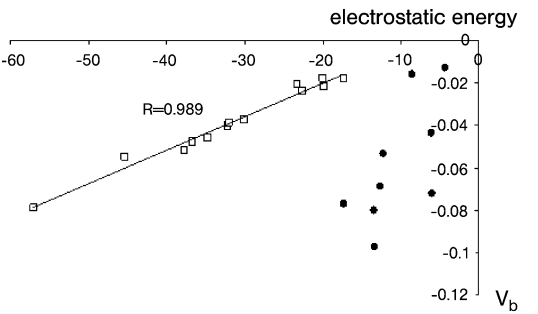


Fig. 17.9 Relationship between the electrostatic energy component (kcal mol^{-1}) and the potential energy density at the H···Y BCP (V_b in au) Circles denote CAHB-DHBs and squares denote RAHBs.

posed to the delocalization and the exchange terms, is not the "driving" interaction for very strong hydrogen bonds.

17.5 Various Other Interactions Related to Hydrogen Bonds

17.5.1 $H^+ \cdots \pi$ Interactions

Figure 17.3 shows the range of dihydrogen bonds, from very strong charge-assisted DHBs to very weak bonds bordering with van der Waals interactions. Interactions such as $\pi\cdots H^+\cdots\pi$ in $C_2H_2\cdots H^+\cdots C_2H_2$ and $C_2H_4\cdots H^+\cdots C_2H_2$ complexes [57], which are related to hydrogen bonding, are also included in the fig-

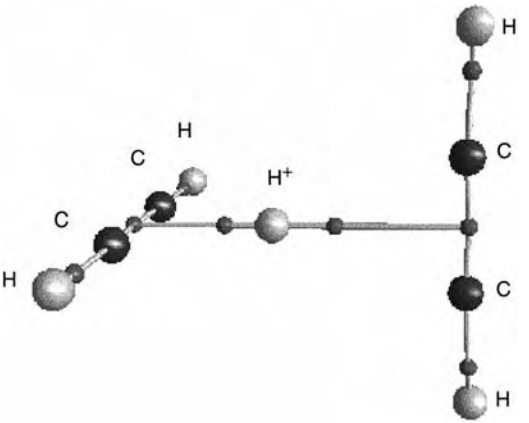


Fig. 17.10 The $H^+\cdots\pi$ interaction for acetylene.

ure. For both complexes the proton is more firmly bound to one of two available π-electron systems. For $C_2H_2\cdots H^+\cdots C_2H_2$, the proton is closer to one of the acetylene molecules (Fig. 17.10); for $C_2H_4\cdots H^+\cdots C_2H_2$ the proton is closer to C_2H_4. In both complexes $C_2H_2\cdots H^+$ and $C_2H_4\cdots H^+$ can therefore be treated as proton-donating systems. For the two short $H^+\cdots\pi$ contacts, the corresponding Laplacians ($\nabla^2\rho_{H+\cdots\pi}$) are negative, indicating that the proton is covalently bonded to the π-electron system. For the longer $H^+\cdots\pi$ contacts, the values of $\nabla^2\rho_{H+\cdots\pi}$ are positive but the corresponding H_b values are negative, indicative of partly covalent nature. Figure 17.3 shows that the $-V_b/G_b$ ratios for such complexes are extremely high, despite the longer $H^+\cdots\pi$ distances, both falling outside the main trend of the relationship.

It has also been found that the delocalization interaction energy is important in these complexes. For the $C_2H_2\cdots H^+\cdots C_2H_2$ complex the electrostatic, exchange, delocalization, and correlation energy terms are -24.5, 46.9, -33.2, and -8.4 kcal mol^{-1}, respectively. For the $C_2H_4\cdots H^+\cdots C_2H_2$ complex these interaction energy terms are -13.1, 19.9, -11.2 and -5.3 kcal mol^{-1}, respectively.

Studies on $\pi\cdots H^+\cdots\pi$ bond complexes indicate these interactions can be regarded as very strong partly covalent hydrogen bonds. The binding energies for these complexes are -19.1 and -9.7 kcal mol^{-1} if the $C_2H_2\cdots H^+$ and $C_2H_4\cdots H^+$ moieties, respectively, are the proton donors to acetylene.

Summarizing, $\pi\cdots H^+\cdots\pi$ bond complexes are characterized by a relatively high $-V_b/G_b$ and $E_{DEL}^{(R)}/E_{EL}^{(1)}$ values which are unexpectedly high if the H$\cdots$Y distance is taken into account (for other hydrogen bonds with the similar H$\cdots$Y distances, these ratios are much smaller). It seems this is a common feature of the interactions between protons and π-electrons. For the T-shaped $FH\cdots C_2H_2$ complex and for the acetylene dimer, in both of which π-electrons are the proton acceptor, the H$\cdots$Y distance (where Y is the middle of a C≡C triple bond) is equal to

2.186 and 2.697 Å, respectively, whereas the corresponding ratio $E_{DEL}{}^{(R)}/E_{EL}{}^{(1)}$ is 0.45 and 0.25, respectively, and the $-V_b/G_b$ ratio is equal to 0.68 and 0.79, respectively.

17.5.2 Hydride Bonds

DHBs may be regarded as protic–hydric interactions, because the protic $X–H^{+\delta}$ bond contains a hydrogen atom carrying a positive charge whereas the hydric hydrogen ($^{-\delta}H–Y$) acts as the proton acceptor [58]. Rozas et al. [58] suggested that $X\cdots{}^{-\delta}H–Y$ systems are also possible in, for example, Li–H···Li and Be–H···Be. These authors named this interaction "inverse hydrogen bonding". It has also been suggested this interaction is called the "hydride bond" [59], because the term "inverse" is usually reserved for so-called blue-shift hydrogen bonds [60]. This type of interaction ($B–H\cdots Na^+$) has been observed in an experimental crystal structure [61]. The hydride bonding in the $BeH_2\cdots Li^+$, $BeH_2\cdots Na^+$, and $BeH_2\cdots Mg^{2+}$ linear complexes, and the variation–perturbation partitioning of interaction energy, have recently been investigated by use of the QTAIM [59]. The binding energies calculated at the MP2/aug-cc-pVQZ level of theory for these systems are −13.3, −11.7, and −59.7 kcal mol^{-1}, respectively. The delocalization energy term is the most important attractive term in these complexes. Similar results were obtained for agostic bonds, delocalization being responsible for stabilization of the systems. It is worth noting that the agostic bonds in $CH_4\cdots LiNH_2$, $CH_4\cdots NaNH_2$, and $CH_4\cdots Na^+$ have weak-to-moderate binding energies ranging from 2 to 6 kcal mol^{-1} [59].

Figure 17.11 shows a classification of the interactions discussed. The arrows indicate electron transfer. One may expect that agostic bonding is a special kind of hydride bonding, because in the former the C–H (or Si–H) bond acts as a Lewis base and metal centre acts as a Lewis acid.

The direction of electronic charge transfer for numerous complexes bound by a variety of interactions (Fig. 17.11) has been analyzed at different levels of approximation. In one study [62] the CHelpG scheme was used to calculate the atomic charges. For the conventional O–H···O hydrogen bond in the *trans*-linear water dimer, the transfer of electronic charge from acceptor to the donor amounts to 32 millielectrons (me) at the MP2/aug-cc-pVTZ//MP2/aug-cc-pVDZ level. More significant charge transfer of 205 me (from BeH_2 to the hydronium ion) is observed in the strong charge-assisted dihydrogen bond of the $H_2OH^+\cdots HBeH$ complex at the MP2/aug-cc-pVDZ level. For hydride bonding in the $HBeH\cdots Li^+$ system electron transfer to Li^+ cation amounts to 41 me at the MP2/aug-cc-pVQZ level and for $CH_4\cdots LiNH_2$, $CH_4\cdots NaNH_2$, and $CH_4\cdots Na^+$ transfer to the metal is 41, 31, and 52 me, respectively, at the MP2/aug-cc-pVDZ level. Only for the strongest ($HBeH\cdots Mg^{2+}$) interaction is H_b negative (at the MP2/aug-cc-pVTZ level) – H_b has positive values for the other complexes. This means that hydride bonds, even when very strong, cannot be classified as a covalent interaction.

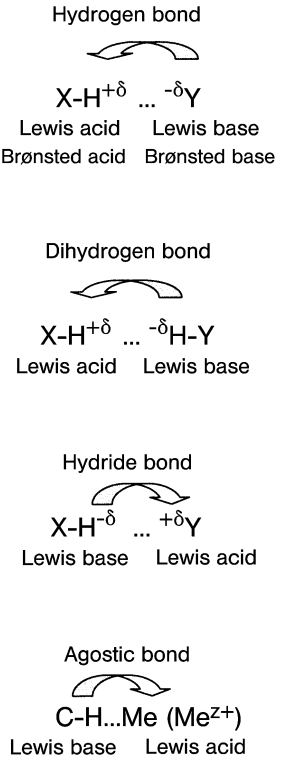


Fig. 17.11 Classification of hydride bonds.

17.6 Summary

A wide range of interactions can be classified as hydrogen bonding. The proton···acceptor distance (H···Y) is an approximate measure of the strength of such interactions. The consensus in the literature is that short H···Y distances imply covalent hydrogen bonding which turns gradually for longer distances into an essentially electrostatic interaction [18, 38, 39], the latter characterization is consistent with Pauling's definition of the hydrogen bond [7]. This also agrees with the QTAIM, because for short H···Y distances the electron density at the BCP is high, similar to that of covalent and polar bonds, and occasionally the Laplacian and/(or at least) H_b is negative. It has also been found that the $-V_b/G_b$ ratio increases with decreasing H···Y distance. A similar relationship has been observed for the ratio of the delocalization and electrostatic interaction energy terms; this ratio increases for shorter proton–acceptor distances, indicative of greater importance of the delocalization energy for shorter distances. The latter relationship is valid for hydrogen bonds and sub-class of these interactions –

dihydrogen bonds. In other words this is unique for Brønsted acid–Brønsted base interactions. For hydride bonds the delocalization interaction energy is the dominant attractive term. Topological data clearly show, however, that such interactions, despite of the significant binding energies, are not covalent, because the values of the Laplacian at the BCPs of Lewis acid–Lewis base contacts are positive. One may, therefore, expect that the increased covalence of hydrogen bonding connected with the increasing importance of the delocalization energy is the unique feature of this interaction.

Finally, DHBs are found to be a subclass of typical hydrogen bonds, ranging from very weak van der Waals-like interactions to very strong bordering on covalent bonds, as is also observed for hydrogen bonds.

References

1 C. F. Matta, J. Hernández-Truillo, T.-H. Tang, R. F. W. Bader, *Chem. Eur. J.* **2003**, 9, 1940–1951.
2 G. A. Jeffrey, W. Saenger, ''Hydrogen Bonding in Biological Structures'', Springer, Berlin **1991**.
3 G. A. Jeffrey, ''An Introduction to Hydrogen Bonding'', Oxford University Press, New York, **1997**.
4 G. R. Desiraju, T. Steiner, ''The weak hydrogen bond in structural chemistry and biology'', Oxford University Press, Inc., New York, **1999**.
5 S. Scheiner, ''Hydrogen Bonding: A Theoretical Perspective'', Oxford University Press: New York, **1997**.
6 R. F. W. Bader, ''Atoms in Molecules. A Quantum Theory'', Oxford University Press, New York, **1990**.
7 L. Pauling, ''The Nature of the Chemical Bond'', Ithaca, NY: Cornell University Press, Ithaca, New York, third edition, **1960**.
8 R. F. W. Bader, *Chem. Rev.* **1991**, 91, 893–928.
9 M. S. Gordon, J. H. Jensen, *Acc. Chem. Res.* **1996**, 29, 536–543.
10 I. Rozas, I. Alkorta, J. Elguero, *J. Am. Chem. Soc.* **2000**, 122, 11154–11161.
11 T. K. Ghanty, V. N. Staroverov, P. R. Koren, E. R. Davidson, *J. Am. Chem. Soc.* **2000**, 122, 1210–1214.
12 E. D. Isacs, A. Shukla, P. M. Platzman, D. R. Hamann, B. Barbiellini, C. A. Tulk, *Phys. Rev. Lett.* **1999**, 82, 600–603.
13 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2005**, 109, 4331–4341.
14 D. Cremer, E. Kraka, *Angew. Chem. Int. Ed. Engl.* **1984**, 23, 627–628.
15 M. Domagała, S. J. Grabowski, *J. Phys. Chem. A* **2005**, 109, 5683–5688.
16 O. Knop, K. N. Rankin, R. J. Boyd, *J. Phys. Chem. A* **2003**, 107, 272–284.
17 K. N. Robertson, O. Knop, T. S. Cameron, *Can. J. Chem.* **2003**, 81, 727–743.
18 G. R. Desiraju, *Acc. Chem. Res.* **2002**, 35, 565–573.
19 R. Parthasarathi, V. Subramanian, N. Sathyamurthy, *J. Phys. Chem. A* **2006**, 110, 3349–3351.
20 S. J. Grabowski, A. J. Sadlej, W. A. Sokalski, J. Leszczynski, *Chem. Phys.* **2006**, 327, 151–158.
21 S. Wojtulewski, S. J. Grabowski, *J. Mol. Struct.* **2002**, 605, 235–240.
22 G. Pimentel, A. McClellan, ''*The hydrogen bond*'', Freeman, San Francisco, **1960**.
23 D. J. Suttor, *J. Chem. Soc.* **1963**, 1105–1110.
24 R. Taylor, O. Kennard, *J. Am. Chem. Soc.* **1982**, 104, 5063–5070.
25 M. A. Viswamitra, R. Radhakrishnan, J. Bandekar, G. R. Desiraju, *J. Am. Chem. Soc.* **1993**, 115, 4868–4869.
26 J. M. A. Robinson, B. M. Kariuki, R. J. Gough, K. D. M. Harris, D. Philp, *J. Solid State Chem.* **1997**, 134, 203–206.

27 S. Harder, *Chem. Eur. J.* **1999**, 5, 1852–1861.

28 S. J. Grabowski, *J. Chem. Res.* **1996**, 534–535.

29 M. Nishio, M. Hirota, Y. Umezawa, ''The CH/π Interaction, Evidence, Nature, and Consequences'', Wiley–VCH, **1998**.

30 T. B. Richardson, S. de Gala, R. H. Crabtree, *J. Am. Chem. Soc.* **1995**, 117, 12875–12876.

31 H. Cybulski, M. Pecul, J. Sadlej, T. Helgaker, *J. Chem. Phys.* **2003**, 119, 5094–5104.

32 I. Rozas, I. Alkorta, J. Elguero, *Chem. Phys. Lett.* **1997**, 275, 423–428.

33 S. J. Grabowski, *J. Phys. Chem. A* **2000**, 104, 5551–5557.

34 S. J. Grabowski, *J. Phys. Chem. A* **2001**, 105, 10739–10746.

35 S. J. Grabowski, T. L. Robinson, J. Leszczynski, *Chem. Phys. Lett.* **2004**, 386, 44–48.

36 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2004**, 108, 5823–5830.

37 E. Espinosa, I. Alkorta, J. Elguero, E. Molins, *J. Chem. Phys.* **2002**, 117, 5529–5542.

38 G. Gilli, F. Bellucci, V. Ferretti, V. Bertolasi, *J. Am. Chem. Soc.* **1989**, 111, 1023–1028.

39 P. Gilli, V. Bertolasi, V. Ferretti, G. Gilli, *J. Am. Chem. Soc.* **1994**, 116, 909–915.

40 G. Gilli, P. Gilli, *J. Mol. Struct.* **2000**, 552, 1–15.

41 H. Umeyama, K. Morokuma, S. Yamabe, *J. Am. Chem. Soc.* **1977**, 99, 330–343.

42 G. Chałasiński, M. M. Szczęśniak, *Chem. Rev.* **1994**, 94, 1723–1765.

43 W. A. Sokalski, S. Roszak, K. Pecul, *Chem. Phys. Lett.* **1988**, 153, 153–159.

44 W. A. Sokalski, S. Roszak, *J. Mol. Struct. (Theochem)* **1991**, 234, 387–400.

45 S. M. Cybulski, G. Chałasiński, R. Moszyński, *J. Chem. Phys.* **1990**, 92, 4357–4363.

46 X. Fradera, M. A. Austen, R. F. W. Bader, *J. Phys. Chem. A* **1999**, 103, 304–314.

47 J. Poater, X. Fradera, M. Solà, M. Duran, S. Simon, *Chem. Phys. Lett.* **2003**, 369, 248–255.

48 S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, 19, 553–561.

49 A. Kaczmarek, A. J. Sadlej, J. Leszczynski, *Mol. Phys.* **2006**, 104, 395–407.

50 M. C. Daza, J. A. Dobado, J. M. Molina, *J. Chem. Phys.* **1999**, 110, 11806–11813.

51 D. Hugas, S. Simon, M. Duran, *Struct. Chem.* **2005**, 16, 257–263.

52 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A* **2006**, 110, 4772–4779.

53 S. J. Grabowski, W. A. Sokalski, E. Dyguda, J. Leszczynski, *J. Phys. Chem. B* **2006**, 110, 6444–6446.

54 E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, 285, 170–173.

55 S. J. Grabowski, W. A. Sokalski, *J. Phys. Org. Chem.* **2005**, 18, 779–784.

56 A. S. Mildvan, M. A. Massiah, T. K. Harris, G. T. Marks, D. H. T. Harrison, C. Viragh, P. M. Reddy, and I. M. Kovach, *J. Mol. Struct.* **2002**, 615, 163–175.

57 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *J. Phys. Chem. A*, **2004**, 108, 1806–1812.

58 I. Rozas, I. Alkorta, J. Elguero, *J. Phys. Chem. A* **1997**, 101, 4236–4244.

59 S. J. Grabowski, W. A. Sokalski, J. Leszczynski, *Chem. Phys. Lett.* **2006**, 422, 334–339.

60 P. Hobza and Z. Havlas, *Chem. Rev.* **2000**, 100, 4253–4264.

61 F. A. Cotton, J. H. Matonic, C. A. Murillo, *J. Am. Chem. Soc.* **1998**, 120, 6047–6052.

62 C. M. Breneman, K. B. Wiberg, *J. Comp. Chem.* **1990**, 11, 361–373.

Part V
Application to Biological Sciences and Drug Design

18
QTAIM in Drug Discovery and Protein Modeling

Nagamani Sukumar and Curt M. Breneman

18.1
QSAR and Drug Discovery

The introduction of a new drug to the market is often the culmination of a long and arduous process of laboratory experimentation, lead-compound discovery, animal testing and preclinical and clinical trials – a process which can typically take as long as 10–15 years from hit to lead to marketable drug. On average, 9 out of 10 promising leads fail, often at an advanced stage in the drug discovery pipeline, because of adverse ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties. One of the most attractive strategies for streamlining and accelerating the process of drug discovery is virtual high-throughput screening (VHTS), employing quantitative structure–activity/property relationships (QSAR/QSPR) modeling. The goal of QSAR/QSPR is the development of correlations between molecular structure and pharmaceutical properties, thereby transforming the search for compounds with specific properties, by use of chemical intuition and experience, into a mathematically quantified and computerized form. When a correlation between structure and activity/property is found and validated, any number of compounds from large pharmaceutical databases, including those not yet synthesized, can be virtually screened on the computer to select structures with the desired properties. Virtual screening using ADMET filters can eliminate compounds likely to have adverse side-effects, identifying the ''losers'' early in the process, to achieve the desired objective of ''fail early, fail cheaply''. The most promising compounds can then be chosen for laboratory synthesis and preclinical testing, thereby conserving resources and accelerating the process of drug discovery.

QSAR and QSPR have proved highly effective within homologous sets of molecules, as is apparent from the extensive literature on the subject [1, 2]. Traditional QSAR methods have not, however, been as successful when applied to more structurally diverse sets of data. This difficulty is partly because of the type of molecular property descriptors used and partly because of the complexity of chemistry space. Descriptors representing simple molecular properties were

often favored in early studies, because they seemed to provide intuitive insight into the physicochemical nature of the biological activity or chemical property under consideration. In recent years, descriptors that correlate with less clearly defined intermolecular interactions have often been found to lead to models with better predictive power [3–11]. Clark [12–14] has argued that the use of descriptors based on local properties calculated at the molecular van der Waals surface, and that do not encode the chemical constitution directly, are likely to provide more generalizable QSPR models that encourage scaffold hopping between diverse regions of chemistry space.

Quantum-chemical descriptors are derived from actual molecular electron density distributions and are readily accessible via ab-initio or semi-empirical calculations. QSAR/QSPR models employing electron-density-derived descriptors are thus applicable to a wide variety of molecules and have the required flexibility to compute physical, chemical and biological properties. The primary disadvantage of such descriptors is the intensive computational effort required to generate them via quantum-chemical calculations, precluding their routine use for large biological molecules or large pharmaceutical datasets. This drawback is circumvented in a QTAIM-based approach, described in Section 18.2. The atom typing scheme and generation of the transferable atom equivalent (TAE) library are outlined in Section 18.3, and TAE reconstruction and descriptor generation are dealt with in Section 18.4. Several families of QTAIM-based descriptors are presented in Section 18.5 and a few sample applications in Section 18.6.

18.2
Electron Density as the Basic Variable

In 1964, Hohenberg and Kohn [15] proved that the external potential $v(\mathbf{r})$ is determined, within a trivial additive constant, by the distribution of electron density $\rho(\mathbf{r})$. Because $\rho(\mathbf{r})$ determines the number of electrons:

$$N = \int \rho(\mathbf{r})\, d\mathbf{r} \tag{1}$$

it follows that $\rho(\mathbf{r})$ also uniquely determines the ground state wave function ψ, the ground state electronic energy, the molecular structure, and all the other electronic properties of the molecule. Thus Bader et al. [16–20] have shown that the topology of the gradient paths of the electron density, $\nabla\rho(\mathbf{r})$, completely specifies the molecular graph. In 1981 Riess and Münch [21] extended the Hohenberg and Kohn theorem to subdomains of a bounded quantum system and showed that the ground-state density of an arbitrary subdomain uniquely determines the ground-state properties of this or any other domain. In fact, any nonzero volume part of the nondegenerate ground-state electron density contains all information about the molecule. This has been termed the holographic electron density theorem [22, 23]. Further, all information about latent molecular properties not

exhibited by a given molecular structure but exhibited by the same molecule in a different state or conformation is fully encoded in any nonzero volume of the nondegenerate ground state electron density. A latent property may be regarded as the response of a molecule to a specific interaction; this principle provides further theoretical justification for QSAR, because most biological activity depends not on the properties of isolated molecules in their equilibrium geometries but on the response of molecules to complex intermolecular interactions.

Although the electron density, in any finite region of space, encodes all molecular properties, in accordance with the Hohenberg–Kohn and Reiss–Münch theorems, the density itself is a rather insensitive function of the atomic and molecular environment, because of the "near-sightedness" of electronic matter [24–26] – i.e. $\rho(\mathbf{r})$ depends significantly only on the potential $v(\mathbf{r}')$ at points $\mathbf{r}'$ near $\mathbf{r}$. The effect on $\rho(\mathbf{r})$ of changes of $v(\mathbf{r}')$ at distant points $\mathbf{r}'$ beyond a cut-off distance R $(\forall|\mathbf{r}'| > R)$ decays monotonically as a function of R. To within an accuracy of $\delta\rho$ the electron density $\rho(\mathbf{r})$ cannot "see" any perturbation beyond the distance R. This "near-sightedness" is, in fact, what makes the study of chemistry more than just an encyclopedic catalog of properties of individual molecules and enables the approximate transferability of atomic and functional group properties from one molecule to another in a similar environment. Understanding the physics and chemistry of large molecules would have been impossible if not for the transferability principle. Because of the Reiss–Münch theorem, perfect transferability of an atom or functional group between molecules is an unachievable limit, although it can be approached very closely [27]. This approximate transferability of fragment properties lies at the heart of the chemical effectiveness of the QTAIM, which is exploited in the transferable atom equivalent reconstruction (TAE RECON) method. The locality principle is also behind other local computational methods, for example the divide and conquer scheme [28, 29]. The relationship between the concepts of transferability and similarity has been discussed by Matta [30] in terms of the short-range nature of the reduced first-order density matrix.

In contrast with other fragment-based electron density reconstruction techniques [31–33], TAE descriptors encode the distributions of electron density-based molecular properties, for example kinetic energy densities [18], local average ionization potentials [34], electrostatic potentials [35–38], Fukui functions [39–42], electron density gradients, and second derivatives or Laplacian distributions [18] (Section 18.3), rather than the density itself. The TAE RECON method is an attractive formulation for rapid molecular electron density reconstruction. QTAIM-based descriptors are capable of generating high-quality models when used with modern machine learning techniques, for example principal-component analysis, artificial neural networks [43–45], kernel partial least squares (KPLS), or support vector machine (SVM) regressions [46], with feature selection accomplished using genetic algorithms [47], sensitivity analysis [48], or a 1-norm linear support vector regression (SVR).

The TAE RECON method is based on the QTAIM [17, 18], wherein an atom in a molecule is defined as the union of an attractor (usually an atomic nucleus) and

its basin (the electron density distribution $\rho(\mathbf{r})$), bounded by an atomic surface of zero flux in the gradient of the electron density:

$$\nabla\rho(\mathbf{r}).\mathbf{n}(\mathbf{r}) = 0, \quad \text{for all } \mathbf{r} \text{ belonging to the surface } S(\Omega) \tag{2}$$

where $\mathbf{n}(\mathbf{r})$ is a unit vector normal to the surface. This is the boundary condition necessary for application of Schwinger's principle of stationary action to an open system [49]. Atoms defined in this way have been shown to satisfy the virial theorem. An extensive body of work [16–19, 49–63] has revealed that virial partitioning gives a natural and rigorous meaning to the intuitive concept of an atom in a molecule, and convergence of the electrostatic interaction based on topological atoms has been computationally demonstrated [64–66]. Atoms defined in this way have uniquely identifiable properties that are approximately additive and transferable from one molecule to another. This transferability feature is really the basis of the TAE method [3, 65–70], because it enables transfer of atomic properties calculated using ab-initio methods in a small molecule to a much larger molecule *containing the same (or very similar) type of atom*. This naturally brings up the question of how to define a TAE atom type, which will be discussed in Section 18.3.

18.3 Atom Typing Scheme and Generation of the Transferable Atom Equivalent (TAE) Library

The quality of molecular TAE descriptors can only be as good as the atom-type representation in the TAE atom type library; it is, therefore, highly desirable to use a representation that reproduces the bonding environment of the atom being modeled as faithfully as possible. Any method that uses atom-based properties to construct and calculate molecular properties must have a consistent atom-typing scheme. For instance, Kier and Hall's electrotopological state (E-state) atoms are selected on the basis of their element identity, valence state, and number of neighboring hydrogen atoms [71] whereas other methods select atom types based on their element type, valence, and connectivity [72]. The atom types in RECON are defined using several criteria (in descending order of priority):

1. element type or atomic number,
2. coordination number (number of other atoms connected to the atom in question),
3. atomic numbers and coordination numbers of the bonded neighbors,
4. size of the ring system, if any, containing the atom, and
5. next-nearest neighbors for mono-coordinate atoms.

RECON employs a sequential fallback procedure which uses the best available representation for each atom (closest match in the TAE library). A requested atom

type is compared with each atom type in the TAE library in succession – by string comparison to entries in a sorted TAE list file, as described in detail by Breneman and coworkers [3, 65, 66] – until the library atom type string with the best match is found. This atom type is then used to model the requested atom in the molecule. The TAE library in our present implementation of RECON [73] contains 915 atom types, but this number is capable of as much expansion as deemed necessary. In a virtual high-throughput application on a large pharmaceutical dataset, one would, in general, expect to find some atom types that are not as well represented in the library as others. The atom types in each molecule are categorized using the Atomtyper algorithm, which also identifies any new atom types encountered that are not satisfactorily represented in the TAE library and must be generated.

The TAE library of topological atomic charge density fragments is constructed in a form that enables rapid retrieval of the fragments and molecular assembly. Associated with each atomic charge density fragment in the TAE library are the coordinates of the bond critical points (BCP) of the atomic charge density (used to translate and reorient the atomic charge density fragments to the molecular coordinate system for molecular electron density reconstruction and visualization, if desired) and additive atomic charge density-based descriptors that encode electronic and structural information relevant to the chemistry of intermolecular interactions. These descriptors are described in detail in Section 18.5.

The generation of the TAE library starts with identification of atom types (using Atomtyper); this is followed by computation of ab-initio molecular wave functions (Hartree–Fock using the 6-31+G(d) theoretical model) [74]. Determination of the topology of the electron density and location of BCP (points along the bond paths connecting pairs of atoms where the electron density reaches its minimum) is performed using the SADDLE program [75]. The TAE atomic surfaces comprise interatomic surfaces defined by the zero-flux condition (2) and the external van der Waals surface of the atom, defined by the $\rho(\mathbf{r}) = 0.002$ e bohr^{-3} isosurface. Although the external surfaces of an atom in an isolated molecule in free space extend out to infinity, in a real interacting molecule a more meaningful boundary is the distance of nonbonded contacts or the van der Waals surface, which has been shown to correlate well with an electron density isosurface [76]. All TAE descriptors are computed on these $\rho(\mathbf{r}) = 0.002$ e bohr^{-3} isosurfaces. The interatomic surfaces are determined, in the PROAIM program [75], by generating a set of steepest descent paths in electron density radiating outward from the BCP. These zero-flux surfaces and the external van der Waals surfaces together form boundaries for integration of the electron-density-derived properties of each atom within a molecule. Electron-density-derived properties (Section 18.5) of the atomic fragments are then computed for each atom in the TAE library, and the van der Waals surface distributions of these properties are encoded in the form of histograms (Fig. 18.1). Finally, an index file listing all atom types in the TAE library is constructed in a form that enables rapid retrieval and atom type matching (as described above) [65, 66].

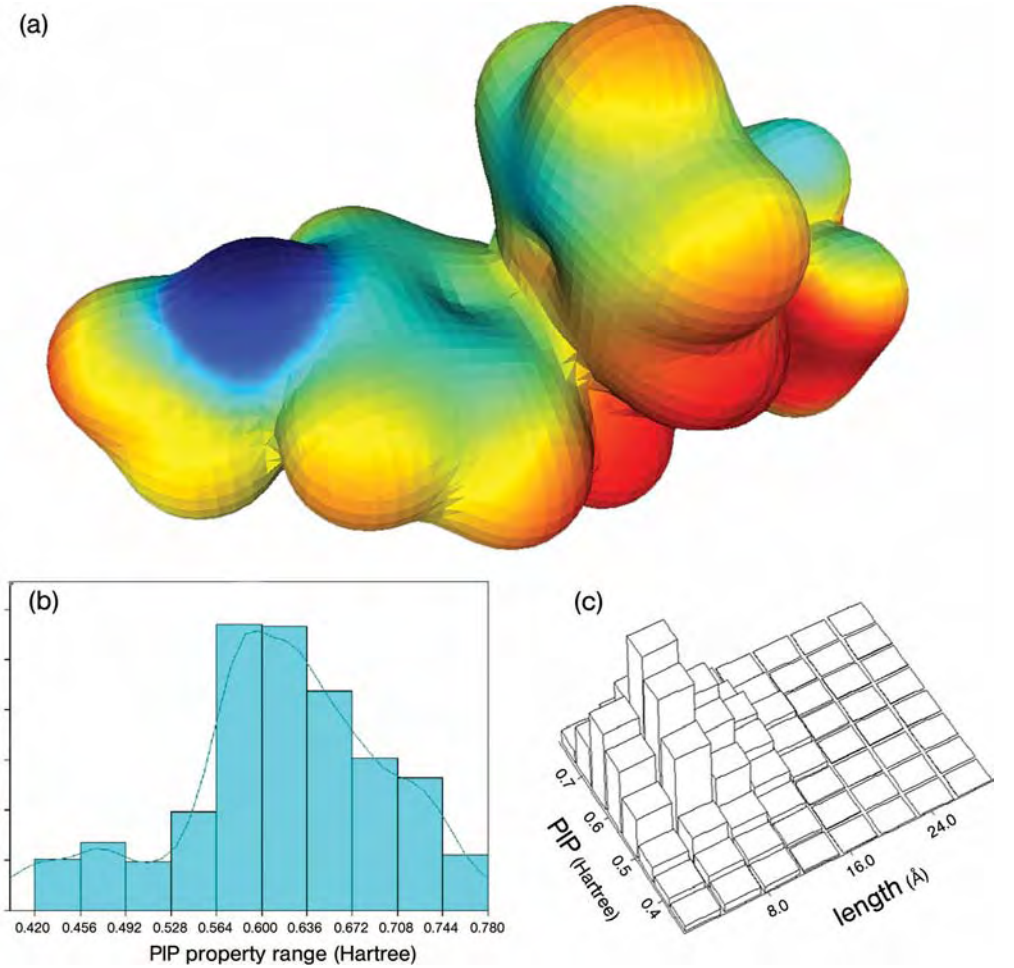


Fig. 18.1 (a) Politzer local average ionization potential PIP(**r**)-encoded van der Waals surface of Pyrilamine, (b) its histogram distribution, and (c) its shape–property histogram distribution from PEST (Adapted from Ref. [132]). The *z*-axis in (b) and (c) is proportional to the surface area of the respective histogram bins.

18.4 TAE Reconstruction and Descriptor Generation

The molecular geometry and/or connectivity information for each input molecule is read at run-time by the RECON algorithm. Atomic connectivity, if not specified in the input, can be determined using a distance criterion – a table of standard single-bond distances is used for this purpose and any pair of atoms with a separation less than 110% of the corresponding single-bond distance are considered

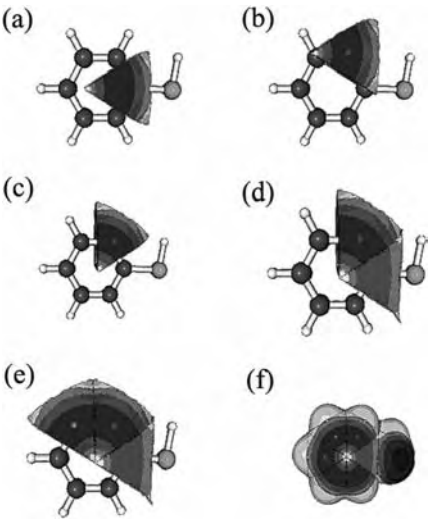


Fig. 18.2 TAE Reconstruction of thiophenol. (a) TAE electron distribution in its native position. In the first step, the first TAE electron-density fragments are translated to the molecular coordinates of the model atom, as shown in (b). The charge-density distribution is then rotated using a quaternion procedure; the results illustrated in (c). These steps are repeated for all atoms in the molecule until the entire molecular charge distribution is reconstructed, as shown in (d–f) [65, 132].

bonded. The size of the ring (if any) that an atom in a molecule belongs to is determined by stepping along the connectivity tree. The current implementation of the RECON algorithm handles connectivity up to four, and rings of three, four, five, and six members are detected, as also are bridgehead atoms, each of which is represented by a distinct atom type.

When atom types and environments have been determined, the closest match is assigned for each atom in the input molecule from the precomputed TAE library of atom types. The densities of the atomic fragments can be combined, if desired, after translating the atomic electron density distributions into the molecular framework and rotating them into the proper orientation by matching the BCP as shown in Fig. 18.2 and described in detail by Whitehead et al. [65].

The molecular TAE descriptors are usually constructed by appropriate arithmetic operations on the respective atomic descriptors stored in the data files that constitute the TAE library. Thus the only computational operations involved in the generation of molecular TAE RECON descriptors are atom-type assignment and matching, then combination of molecular TAE descriptors. This makes the method highly suitable for VHTS applications, because it scales well with both molecular size and the size of the database – a database of 42,689 drug-like molecules from the NCI HIV database could be screened within 7.6 min. on a 1.7 GHz Intel Pentium with 529 KB RAM under the Mandrake Linux operating system (Table 18.1) [77]. The TAE descriptors are described in detail in Section 18.5.

Table 18.1 CPU times in seconds for *RECON* running on an SGI 300 MHz MIPS R12000 processor with FPU 640 MB RAM and IRIX64 release 6.5 and on 1.7 GHz Intel Pentium with 529 KB RAM under the Mandrake Linux operating system [77].

Test dataset	Number of molecules	File format	SGI 300 MHz Octane MIPS R12000 FPU; 640 MB IRIX64		1.7 GHz Intel Pentium Linux; 529 MB RAM	
			User CPU (s)	System CPU (s)	User CPU (s)	System CPU (s)
MAO inhibitors	1650	SDF	102.7	44.5	15.3	3.5
	1641	SMILES	122.3	45.9	61.3	3.6
Proteins	25	PDB	186.8	194.5	65.1	17.6
NCI AIDS	42,689	SDF	2327.2	1131.0	391.0	67.5

18.5 QTAIM-based Descriptors

When all the atoms in a molecule have been typed, the molecular TAE descriptors are computed, usually by simple addition of the corresponding (precomputed) descriptors of the atomic fragments from the TAE library. Some electronic properties retrieved from the library after TAE reconstruction are listed in Table 18.2. These descriptors fall into four general classes:

- traditional descriptors, for example molecular volume, surface area, and dipole moments, computed from the TAE;
- topological descriptors, which depend only on the molecular connectivity, for example the molecular connectivity index (χ_0) [78] and atom type counts;
- electron-density-derived TAE surface descriptors – extrema, surface integral averages, and histogram bins are generated for each of the properties in Table 18.2; and
- descriptors sensitive to the molecular coordinates and requiring a 3D structure for their evaluation, for example RECON autocorrelation descriptors (RAD), PEST shape–property hybrid descriptors [79], based on Zauhar's shape signature ray-tracing scheme [80, 81], and an implementation of GETAWAY descriptors [8, 9] computed from the spatial coordinates of the atoms and based on a leverage matrix, the so-called molecular influence matrix.

Table 18.2 Electron-density-derived properties after molecular reconstruction.

Integrated	Energy	
	Integrated electron population	
	Volume	
	Surface area	
Topological	Molecular connectivity index (χ_0)	
	Topological autocorrelations (TRAD)	

Surface electronic properties – surface extrema, surface integral averages, histogram bins, wavelet coefficients derived from surface distributions, and autocorrelations based on atomic integral averages are available for each property

EP	Electrostatic potential	$EP(\mathbf{r}) = \sum_\alpha \frac{Z_\alpha}{\lvert\mathbf{r}-\mathbf{R}_\alpha\rvert} - \int \frac{\rho(\mathbf{r}')\,dr'}{\lvert\mathbf{r}-\mathbf{r}'\rvert}$
DRN	Electron density gradient normal to 0.002 e bohr^{-3} electron density isosurface	$\nabla\rho$.n
G	Electronic gradient kinetic energy density	$G(\mathbf{r}) = -(1/2)(\nabla\psi^*.\nabla\psi)$
K	Electronic Schrödinger kinetic energy density	$K(\mathbf{r}) = -(1/2)(\psi^*\nabla^2\psi + \psi\nabla^2\psi^*)$
DGN	Gradient of the Schrödinger kinetic energy density normal to surface	∇K.n
DGN	Gradient of the gradient kinetic energy density normal to surface	∇G.n
F	Fukui F^- function scalar value	$F^-(\mathbf{r}) = \left[\frac{\partial\rho(\mathbf{r})}{\partial N}\right]_v \approx \rho_{HOMO}(\mathbf{r})$
L	Laplacian of the electron density	$L(\mathbf{r}) = -\frac{1}{4}\nabla^2\rho(\mathbf{r}) = K(\mathbf{r}) - G(\mathbf{r})$
BNP	Bare nuclear potential	$BNP(\mathbf{r}) = \sum_\alpha \frac{Z_\alpha}{\lvert\mathbf{r}-\mathbf{R}_\alpha\rvert}$
PIP	Local average ionization potential	$PIP(\mathbf{r}) = \sum_i \frac{\rho_i(\mathbf{r})\lvert\varepsilon_i\rvert}{\rho(\mathbf{r})}$

Descriptors requiring 3D coordinates:

RAD	Recon autocorrelation descriptors for all TAE surface properties above
PEST	Shape–property hybrid descriptors for all TAE surface properties above
GETAWAY [8, 9]	Based on a leverage matrix – the molecular influence matrix

The TAE volume and surface area descriptors have meanings similar to those of other volume and surface descriptors computed using most modern molecular modeling programs. Molecular volume is most often associated with hydrophobic effects and tends to be correlated with the energy required to "dig a hole" in the solvent medium for the molecule. This is the sum of the energies required to break existing noncovalent interactions between solvent molecules, and the desolvation energies of the binding site with which the molecule might interact. In the case of solution binding and molecular recognition the desolvation energy of the solute molecule is also related to its volume. Other descriptors computed in

RECON are the total TAE energy, the integrated electron population, and simple topological descriptors, for example the molecular connectivity index and atom type counts [78, 82–91].

18.5.1 TAE Descriptors

Among the TAE descriptors, the gradient of the electron density normal to the molecular surface ($\nabla\rho.\mathbf{n}$), has been used to distinguish "soft" regions of polarizable electron density from more tightly held regions. For example, values of $\nabla\rho.\mathbf{n}$ are much smaller over electron-rich systems and aromatic rings than over polarized or electron-deficient alkyl carbon atoms. Because $\rho(\mathbf{r})$ decreases away from the attractors, $\nabla\rho.\mathbf{n}$ is always negative; large negative values of $\nabla\rho.\mathbf{n}$ indicate the electron density of the underlying molecular region is more tightly held and less likely to extend very far from the molecule.

Histogram bins of the electrostatic potential (EP), defined as:

$$EP(\mathbf{r}) = \sum_{\alpha} \frac{Z_\alpha}{|\mathbf{r} - \mathbf{R}_\alpha|} - \int \frac{\rho(\mathbf{r}')\, d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \tag{3}$$

its surface integral (SIEP), extrema, and integral average (SIEPIA), represent the scalar electrostatic potential values on the surface of the atom or molecule. Electrostatic potential has been implicated in many molecular and intermolecular phenomena, including acid–base interactions, solvation behavior, and pK_a correlations [36, 37, 92]. These descriptors are often found in the best models of hydrogen-bonding systems and in regressions involving polar or dipolar molecules. Donor/acceptor behavior is also modeled well by use of these histogram descriptors. Whereas electropositive and electronegative regions of the molecular surface are represented by the high and low histogram bins of EP, respectively, the middle bins correspond to hydrophobic regions.

Another set of descriptors is derived from the bare nuclear potential (BNP), mapped on to the molecular electron density isosurface. BNP is simply the first term in Eq. (3). Because the geometry and orientation of the nuclei reflect the electron-density distribution, the strength of the BNP field mapped on to an electron density isosurface indicates the regions of imbalance between the nuclear–electron attractive forces and the interelectron repulsive forces. This provides information complementary to that obtained via the electrostatic potential.

The electronic gradient kinetic energy density distribution, G is defined as:

$$G(\mathbf{r}) = -\nabla\psi^*.\nabla\psi \tag{4}$$

The surface integral (SIG) of G can be interpreted as being associated with differences in donor/acceptor activity. The Schrödinger kinetic energy density, K, given by:

$$K(\mathbf{r}) = -(\psi^*\nabla^2\psi + \psi\nabla^2\psi^*) \tag{5}$$

is a rather smooth function over the surface of a typical molecule and is most negative in those portions of space where there is a local concentration of negative charge [18]. This also corresponds to areas of negative Laplacian, because imbalances of K and G are responsible for nonzero Laplacian values:

$$L(\mathbf{r}) = -\frac{1}{4}\nabla^2\rho(\mathbf{r}) = K(\mathbf{r}) - G(\mathbf{r}) \tag{6}$$

The Laplacian ($\nabla^2\rho$) is the trace of the second derivative matrix of the electron density at any point in space and has been extensively studied by Bader et al. It has been implicated as a descriptor in the selectivity of electrophilic aromatic substitution and in donor–acceptor interactions. The truly indicative regions are the negative Laplacian peaks, which form near the outer core regions of the electron density of molecules. When these peaks are in regions of nonbonded electron density or in regions of electrophilic attack their magnitude tends to be useful for prediction of the rates of these reactions [18, 93]. Such ''negative Laplacian peaks'' are usually seen within 0.25–0.4 Å of an electron donor atom – well within the molecular van der Waals surface chosen for this analysis. ''Shadows'' of these internal extrema are, nevertheless, often present on the molecular surface. These ''Laplacian shadows'' on the molecular van der Waals surface are the best indicators of what is going on inside the molecular surface. These surface manifestations of the internal Laplacian peaks are often opposite in sign to that of the actual peak, as a result of Laplacian normalization. Consequently, slightly less negative regions of surface values of K often indicate the presence of Brönsted bases [94].

The rate of change of the K electronic kinetic energy density normal to and away from the molecular surface ($\nabla K.\mathbf{n}$) has been shown to describe differences between the polarizability and hydrophobicity of molecular regions [94]. More negative ranges of this function indicate that the region is more hydrophobic and less susceptible to electrophilic attack [18, 59, 95]. Likewise, $\nabla G.\mathbf{n}$ and $\nabla\rho.\mathbf{n}$ are often significant in correlation models of dispersion interactions [94].

One of the most interesting and underexploited descriptors is the local average ionization potential, called ''I-bar'', of the GIPF parameter set of Politzer et al. [34, 96], referred to here as the Politzer ionization potential (PIP):

$$PIP(\mathbf{r}) = \sum_i \frac{\rho_i(\mathbf{r})|\varepsilon_i|}{\rho(\mathbf{r})} \tag{7}$$

PIP histogram descriptors appear in many diverse models of disparate phenomena. PIP is correlated with several intermolecular binding modes, for example induced-dipole interaction. PIP descriptors also carry information about the ''hardness'' or ''softness'' of a region of electron density, and donor–acceptor information. Quite frequently, PIP descriptors occur, with $\nabla\rho.\mathbf{n}$ and SIK, in models describing differential solubility or hydrophobic–hydrophilic interaction tendency.

Another class of molecular and regional descriptors is derived from the Fukui radical reactivity indices (F), defined as:

$$F(\mathbf{r}) = \left[\frac{\partial\rho(\mathbf{r})}{\partial N}\right]_v \approx \frac{1}{2}(\rho_{LUMO}(\mathbf{r}) + \rho_{HOMO}(\mathbf{r})) \tag{8}$$

The Fukui index defined above describes radical reactivity. Similar Fukui indices:

$$F^- \approx \rho_{\mathrm{HOMO}}(\mathbf{r}) \tag{9}$$

and

$$F^+ \approx \rho_{\mathrm{LUMO}}(\mathbf{r}) \tag{10}$$

describe reactivity toward electrophilic and nucleophilic attack, respectively. The Fukui indices are somewhat related to PIP, in that both involve a perturbation expression which is meant to describe the spatial distribution of radical reactivity. For PIP the molecular surface is encoded with energy-weighted orbital densities whereas for F there is a selectable denominator term which places the reactivity index on a cationic, radical, or anionic scale.

Other electron-density-derived descriptors used in the literature [13, 14] include the local electron affinity:

$$\mathrm{EA}(\mathbf{r}) = \frac{-\sum_{i=\mathrm{LUMO}}^{N} \rho_i(\mathbf{r})\varepsilon_i(\mathbf{r})}{\sum_{i=\mathrm{LUMO}}^{N} \rho_i(\mathbf{r})} \tag{11}$$

the electronegativity:

$$\chi(\mathbf{r}) = (\mathrm{PIP} + \mathrm{EA})/2 \tag{12}$$

the local hardness:

$$\eta(\mathbf{r}) = (\mathrm{PIP} - \mathrm{EA})/2 \tag{13}$$

and the local polarizability:

$$\alpha(\mathbf{r}) = \frac{-\sum_{i=1}^{N} \rho_i'(\mathbf{r})q_i\bar{\alpha}_i(\mathbf{r})}{\sum_{i=1}^{N} \rho_i'(\mathbf{r})q_i} \tag{14}$$

defined within the framework of semi-empirical MO theory. Ehresmann [13, 14] found that the local electron affinity, local hardness, and local polarizability had little correlation with other descriptors in common use and these descriptors effectively extend the variance of the descriptor set. Use of local electron-density-derived descriptors potentially leads to an increased likelihood of scaffold hopping

(i.e. switching from one structural type to another) in QSAR and virtual screening applications and to more robust and general QSPR models.

Encoding of surface property distributions may be accomplished by use of multiple methods, and use of alternatives to histogram-based representations have often proved useful. One such alternative scheme involves the use of wavelet coefficients to capture TAE-encoded surface-property distributions [97, 98]. In recent years, wavelet encoding has gained popularity in diverse applications as an efficient means of data compression and pattern recognition [97]. The wavelet basis has advantages over the Fourier basis in that, although the trigonometric functions used in Fourier expansion are monochromatic in frequency but entirely delocalized in position, the wavelet basis is well localized in both frequency and position. Wavelet encoding and decoding can be accomplished by use of a simple scaling and dilation algorithm. Wavelet encoding enables a more compact representation of molecular surface property distributions than use of histogram descriptors. Our implementation of TAE property modeling with wavelet coefficient descriptors (WCD) has been described by Sundling et al. [97, 98]. Molecular WCDs can be reconstructed additively from the constituent atomic WCDs.

18.5.2
RECON Autocorrelation Descriptors

RECON autocorrelation descriptors are patterned after whole surface autocorrelation descriptors [79] and are a computationally inexpensive way of including 3D shape information within the TAE RECON formalism. RADs use integrated TAE surface properties (P_x) to calculate property autocorrelations using Gasteiger's formula [99]:

$$A(R_{xy}) = \frac{1}{n}\sum_{x,y} P_x \times P_y \tag{15}$$

The autocorrelation function is then binned by the distance (R_{xy}) between atoms x and y; n is the number of atomic pairs. The electron-density-derived TAE properties are integrated over the external atomic surfaces to compute RAD. Generation of RAD data involves a mere 3–5% CPU overhead over the computation of 2D TAE descriptors for drug-sized molecules. Where 3D structures are not available, the topological distances between pairs of atoms may be used for binning the autocorrelation function, to yield conformation-insensitive 2D RAD or TRAD (topological RECON autocorrelation descriptors).

18.5.3
PEST Shape–Property Hybrid Descriptors

TAE descriptors can be supplemented, with some increase in computation time, by hybrid shape–property descriptors that encode detailed information about molecular shape without requiring an alignment procedure. The supplemental infor-

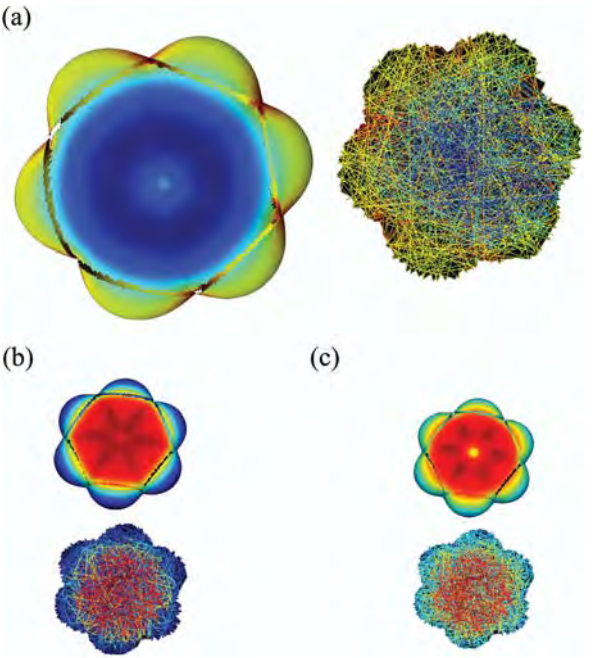


Fig. 18.3 (a) Politzer local average ionization potential PIP(**r**), (b) kinetic energy density $K(\mathbf{r})$, and (c) Laplacian distribution $\nabla^2\rho(\mathbf{r})$ of benzene reconstructed from atomic densities. The seams show the zero-flux surfaces $\nabla\rho.\mathbf{n} = 0$. Shown alongside are the corresponding PEST ray-trace descriptors [79].

mation available from these descriptors is useful where the shape of the molecule plays a determining role in binding. Property-encoded surface translator (PEST) descriptors [79] may be computed using atomic fragment-based TAE RECON property-encoded surface reconstructions, as shown in Fig. 18.3 (or using ab-initio or semi-empirical electron-density surfaces and electronic properties). The Zauhar shape signature ray-tracing scheme [80, 81], upon which PEST descriptors are based, seeks to encode the shape of the molecular volume by using the distribution of ray lengths obtained by performing a ray-tracing procedure within the molecular van der Waals envelope, beginning from an arbitrary starting position. The converged ray-length distribution then represents a distinctive "shape signature" of the molecule. PEST records the ray lengths and TAE properties at each point where the rays encounter the molecular surface, to generate two-dimensional hybrid shape–property histogram descriptors, as shown in Fig. 18.1c. The algorithm for combining the densities of the TAE fragments, after translating them into the molecular framework and rotating them into the proper orientation by matching up BCP, is described by Whitehead et al. [65] and the ray-tracing and descriptor computation algorithms are described by Breneman et al. [79]. Inclusion of PEST hybrid shape–property descriptors with 2D topolog-

ical descriptors increases the predictive capability of QSAR and QSPR models. The information content of PEST shape–property descriptors has been shown to be comparable with or greater than 3D field-based methods such as CoMFA [100], with the considerable advantage of not requiring an explicit alignment rule.

18.5.4
Electron Density-based Molecular Similarity Analysis

Electron density functions have been used to develop the idea of molecular quantum similarity measures (MQSM) for rational drug design and have been extensively applied to pharmacological and toxicological problems [101–115]. MQSM encapsulates the principle that the more similar two molecules are, the more similar will their properties be [101, 116, 117]. The degree of similarity can be established on the basis of the electronic distribution, the topology of the BCP [111–115], or any electron-density-derived property. Vercauteren et al. [118] have developed a procedure for similarity searching of molecules on the basis of comparison of critical point representations of 3D electron-density maps. Pair-wise and multiple comparisons between the molecular critical point graphs are performed using a Monte Carlo/simulated annealing technique. The method has been used for similarity searching of pharmaceutical ligands at different levels of crystallographic resolution.

Girones et al. [119] have developed a kinetic energy density-based molecular quantum similarity measure (KE MQSM) for two quantum objects:

$$Z_{AB} = \int K_A(\mathbf{r}) K_B(\mathbf{r})\, d\mathbf{r} \tag{16}$$

which can be used to construct a Carbo index of similarity:

$$C_{AB} = Z_{AB}(Z_{AA}Z_{BB})^{-1/2} \tag{17}$$

As a means of interpreting and visualizing molecular structure, Ponec [109, 110] introduced the so-called domain-averaged Fermi hole:

$$g_\Omega{}^A(\mathbf{r}_1) = N_\Omega \rho_A(\mathbf{r}_1) - 2\int_\Omega \rho_A(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_2 \tag{18}$$

where

$$N_\Omega = \int_\Omega \rho_A(\mathbf{r})\, d\mathbf{r} \tag{19}$$

is the mean number of electrons in the domain Ω and $\rho_A(\mathbf{r}_1, \mathbf{r}_2)$ is the pair density. Girones and Ponec [120] used the domain-averaged Fermi hole density to define the fragment molecular quantum self-similarity measure for fragment A:

$$Z_{AA}{}^{\Omega} = \int_{\Omega} g_{\Omega}{}^{A}(\mathbf{r}) g_{\Omega}{}^{A}(\mathbf{r})\, d\mathbf{r} \tag{20}$$

Similarity indices based on the density can give exaggerated weight to small mismatches in regions of space with high electron density, e.g. in the vicinity of nuclei, when comparing molecules with slightly different geometry. To provide a more balanced measure of similarity that reflects the reactivity of the molecule without being biased by small nuclear cusp mismatches, Matta [30] proposed the use of the integral of the Laplacian to define a ''reactivity'' similarity index for a pharmacophore:

$$R_{\Omega \cup \Omega'} = \frac{\int_{\substack{\Omega \cup \Omega' \\ \nabla^2\rho(r) \geq 0}} \nabla^2\rho_{\Omega} \nabla^2\rho_{\Omega'}\, dv}{\sqrt{\int_{\substack{\Omega \\ \nabla^2\rho(r) \geq 0}} \nabla^2\rho_{\Omega'}\, dv \int_{\substack{\Omega' \\ \nabla^2\rho(r) \geq 0}} \nabla^2\rho_{\Omega}\, dv}} \tag{21}$$

A recent application of TAE fragments for similarity searching consists of using the statistics of TAE atom type fragments, clustered in accordance with the priority scheme described in Section 18.3, to sample neighborhoods in molecular-property space and to assess the predictivity of models constructed using other descriptors. This assessment is then used to supplement a training set with more molecules in regions of molecular-property space that are poorly represented in the training set. Application to the design of novel selective displacers for protein chromatography has been discussed [121].

In collaboration with our group, Oloff, et al. [122] have developed a novel structure-based cheminformatics approach (CoLiBRI) using TAE RECON descriptors to search for complementary ligands, based on representation of both receptor-binding sites and their respective ligands in a space of universal chemical descriptors. The binding site atoms involved in the interaction with ligands were identified by applying Delaunay tessellation to the X-ray structures of the ligand–receptor complexes. TAE RECON descriptors were calculated independently for each ligand and for its active-site atoms. This representation of both ligands and active sites using the same set of chemical descriptors enables correlation of chemical similarities between active sites and their respective ligands to be elucidated. A procedure for mapping patterns of nearest-neighbor active site vectors in a TAE RECON space on to those of their complementary ligands enables prediction of a virtual complementary ligand vector in the ligand chemical space from the position of a known active site vector; this is followed by estimation of chemical similarity of the virtual ligand vector and molecules in a chemical database, to identify real compounds most similar to the virtual ligand. Thus, knowledge of the structure of the receptor active site enables identification of its complementary ligands in large databases of chemical compounds by use of rapid chemical similarity searches. Conversely, starting from the chemical structure of the ligand one may identify possible complementary receptor cavities. Applied to a data set of 800 X-ray characterized ligand–receptor complexes, knowledge of the active site structure enabled identification of its complementary ligand among the

top 1% of a large chemical database in over 90% of all test active sites when a binding site of the same protein family was present in the training set. When test receptors were highly dissimilar and not present among the receptor families in the training set, CoLiBRI was still able to quickly eliminate 75% of the chemical database as improbable ligands.

18.6 Sample Applications

18.6.1 QSAR/QSPR with TAE Descriptors

Several illustrative applications of QSAR/QSPR modeling with TAE descriptors are shown in Figs 18.4–18.7. Figure 18.4 shows the results of modeling the acute toxicity of organic compounds in fish using RECON descriptors and KPLS models constructed using Analyze [123]. The results are averaged over 100 bootstraps. The dataset [124] comprises 375 molecules, 300 of which are used for training and 75 for testing the model predictions. Predictions on the test set have a

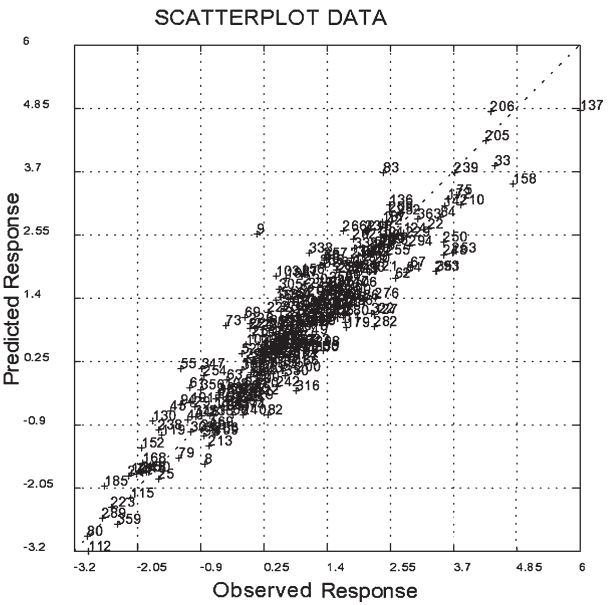


Fig. 18.4 Results from KPLS models for the acute toxicity of 375 organic compounds to fathead minnows [124], constructed with Analyze [123] and RECON descriptors, averaged over 100 bootstraps. $R^2 = 0.86$ for the training set (300 molecules); 0.81 for the test set (75 molecules); leave one out (LOO) $Q^2 = 0.76$.

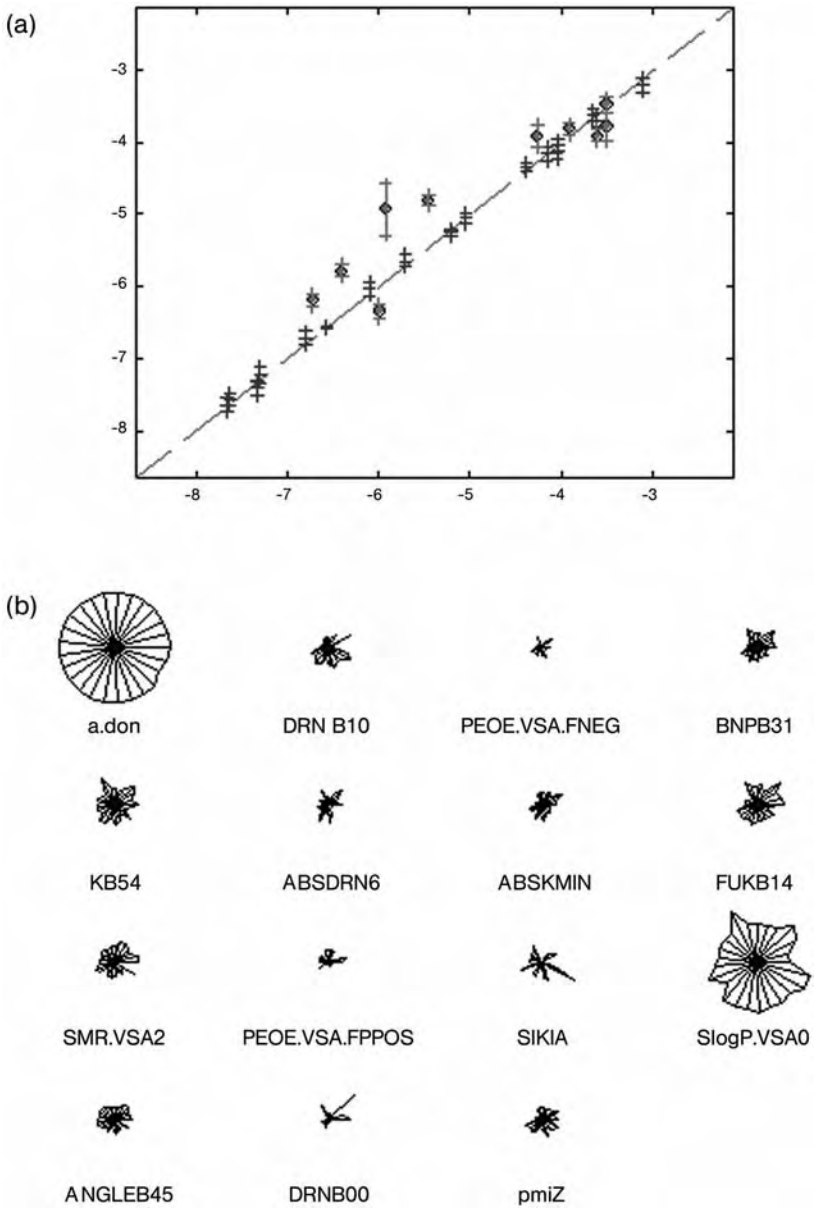


Fig. 18.5 (a) Bagged SVM model for Caco-2 permeability with RECON, MOE, and PEST descriptors using fifteen features. (b) Star plot showing descriptor importance in 20 SVM bootstraps for Caco-2 permeability. The eight descriptors on the left are negatively weighted; the seven on the right are positively weighted; each ray represents a separate bootstrap; the radius of each ray represents the weight or importance of that descriptor in that bootstrap [133].

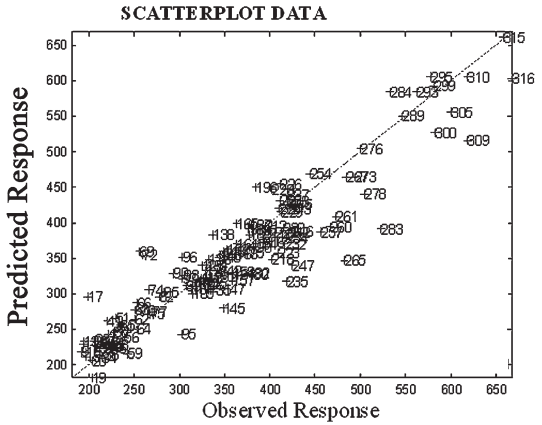


Fig. 18.6 Prediction of the glass transition temperatures of polymers [126] from the Bicerano data set [127] of 300 polymers, 173 in the training set and 127 in the test set. Kernel PLS modeling using TAE/RECON descriptors from repeat unit end-capped with two monomer units, five latent variables. Test set $Q^2 = 0.928$.

Q^2 statistic of 0.81. In this case addition of 2D descriptors does not improve the models.

Figure 18.5 shows the results from a bagged SVM model for Caco-2 permeability with RECON, MOE, and PEST descriptors using fifteen features. Descriptor importance in 20 SVM bootstraps is depicted in the star plot in Fig. 18.5b, in which each ray represents a separate bootstrap and the radius of each ray represents the weight or importance of that descriptor in that bootstrap.

TAE RECON descriptors have also been used for the prediction of polymer properties [125, 126]. Figure 18.6 shows the results from KPLS modeling [126] of the glass transition temperatures of 300 polymers from the Bicerano data set [127], using TAE RECON descriptors of the repeat unit end-capped with two monomer units. One-hundred and seventy-three polymers were used in the training set and 127 in the test set. This model yielded excellent prediction results, as is apparent from the score of $Q^2 = 0.928$ for the test set.

18.6.2
Protein Modeling with TAE Descriptors

Although a limited set of ab-initio computations of some small proteins have recently been reported, routine ab-initio computations on proteins within a reasonable time are not currently feasible. This is where the advantages of the TAE reconstruction method are evident. Two-dimensional TAE descriptors for proteins based on the primary structure, i.e. the amino acid sequence, can be computed very rapidly with RECON, using the amino acid residues as the TAE fragments. Such studies have been performed and used to model the retention times of

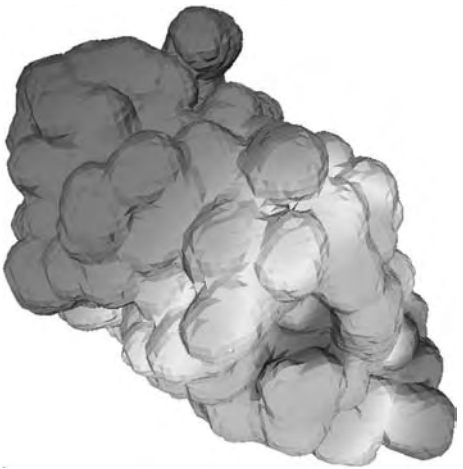


Fig. 18.7 The EP surface of the turkey egg-white lysozyme 135L, computed using RECON through hexadecapole order. The surface here is constructed as a simple union of atom-centered spheres at predetermined van der Waals radii for each atom.

proteins in chromatographic columns under a variety of experimental conditions [128–130]. TAE multipoles can also be used to generate more refined EP surfaces for proteins. As an example, Fig. 18.7 shows the EP surface of a small protein, turkey egg-white lysozyme 135L, computed using RECON through hexadecapole order. Computation of 3D RAD and PEST descriptors using the 3D protein structure and integrated atomic TAE properties entails a small computational overhead but is still accessible even when using very modest computational resources [131].

18.7 Conclusions

We have seen that the QTAIM is a powerful method for rapid generation of ab-initio quality electron densities and electronic properties of large molecules, proteins, and entire of pharmaceutical databases, using TAE RECON technology. Electron-density-derived descriptors generated using the TAE RECON method have also been successfully used to predict diverse molecular properties. Such conformation-insensitive descriptors are valuable in their own right for several reasons:

- For a molecule that has not yet been synthesized, conformational information is not available and can be theoretically computed only by energy minimization, which is often computation-intensive and sensitive to details of the

algorithm and the values used. It is desirable to have a set of theoretical descriptors that have uniquely defined values, irrespective of the details of minimization.

- Even for naturally occurring biomolecules, the active conformation responsible for a certain physiological effect is often unknown. For large fluxional molecules like proteins the minimum energy conformation, even if known, need not correlate with the physiologically active conformation. Further, no single conformation may be sufficient to explain the observed activity.
- RECON descriptors can be supplemented, with some increase in computation time, by hybrid shape–property descriptors from the PEST algorithm. PEST descriptors encode information about the molecular shape, without requiring an alignment procedure for their computation. The supplementary information available from PEST descriptors is useful when the shape of the molecule plays a determining role in binding.
- RAD descriptors may be used, where necessary, to achieve an optimum compromise between conformation sensitivity and computation speed. Computation of RAD incorporates conformational information into the descriptors, while introducing a mere 3–5% CPU overhead.

In this respect and in terms of speed and high-throughput capability, RECON descriptors scale computationally in much the same way as topological descriptors [78, 82–91], with the added advantage that they contain information derived from the electron density distribution, well beyond that contained in indices derived from connectivity alone. The TAE RECON technology has promise for virtual high-throughput screening and design of focused libraries.

Several other electron-density-derived properties are amenable to computational evaluation in a QTAIM framework and are likely to be of value as molecular descriptors in different contexts. Electron-density-derived measures of molecular similarity provide a convenient and fundamental means of exploring chemistry space and assessing the similarities of molecules, the predictive capability of a QSAR model, or the diversity of a molecular library. The QTAIM is also, currently, one of the most computationally tractable methods for going beyond classical models for study of biological molecules and their interactions. CoLiBRI affords rapid prefiltering of large chemical databases to eliminate compounds that have little chance of binding to a receptor active site. Knowledge of the receptor active site structure affords straightforward and efficient identification of its complementary ligands with CoLiBRI; conversely, starting from the ligand chemical structure, one may also identify possible complementary receptor cavities. The QTAIM thus provides the basis for a versatile collection of promising techniques for drug-design applications.

References

1 C. Hansch and A. Leo, *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, ed. S.R. Heller. 1995: American Chemical Society. 1–557.
2 C. Hansch and A. Leo, *Exploring QSAR*. 1995: American Chemical Society.
3 C.M. Breneman and M. Rhem, J. Comp. Chem., 1997. **18**(2): p. 182–197.
4 R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*. Methods and Principles in Medicinal Chemistry, ed. R. Mannhold, H. Kubinyi, and H. Timmerman. Vol. 11. 2000: Wiley–VCH.
5 R.S. Pearlman and K.M. Smith, *Novel Software Tools for Chemical Diversity*, in *3D QSAR and Drug Design: Recent Advances*, H. Kubinyi, Y. Martin, and G. Folkers, Editors. 1997, Kluwer Academic Publishers: Dordrecht, Netherlands.
6 R.S. Pearlman, *Novel Software Tools for Addressing Chemical Diversity* http://www.netsci.org/Science/Combichem/feature08.html
7 R.S. Pearlman and K.M. Smith, Perspectives in Drug Discovery Design, 1998. **9–11**: p. 339–353.
8 V. Consonni, R. Todeschini, and M. Pavan, J. Chem. Inf. Comput. Sci., 2002. **42**: p. 682–692.
9 V. Consonni, R. Todeschini, M. Pavan, and P. Gramatica, J. Chem. Inf. Comput. Sci., 2002. **42**: p. 693–705.
10 J.H. Schuur, P. Selzer, and J. Gasteiger, J. Chem. Inf. Comput. Sci., 1996. **36**: p. 334–344.
11 P. Labute, J. Mol. Graph. Model., 2000. **18**: p. 464–477.
12 T. Clark, J. Molec. Graph. Model., 2004. **22**: p. 519–525.
13 B. Ehresmann, B. Martin, A.H.C. Horn, and T. Clark, J. Mol. Model., 2003. **9**: p. 342–347.
14 B. Ehresmann, M.J.D. Groot, A. Alex, and T. Clark, J. Chem. Inf. Comput. Sci., 2004. **44**(2): p. 658–668.
15 P. Hohenberg and W. Kohn, Phys. Rev., 1964. **136**: p. B864–B871.
16 R.F.W. Bader, S.G. Anderson, and A.J. Duke, J. Amer. Chem. Soc., 1979. **101**(6): p. 1389–1395.
17 R.F.W. Bader, J. Chem. Phys., 1980. **73**(6): p. 2871–2883.
18 R.F.W. Bader, *Atoms in Molecules: A Quantum Theory*. 1990, Oxford: Oxford Press.
19 R.F.W. Bader, Chem. Rev., 1991. **91**(5): p. 893–928.
20 R.F.W. Bader, J. Phys. Chem. A, 1998. **102**(37): p. 7314–7323.
21 I. Riess and W. Münch, Theor. Chim. Acta, 1981. **58**: p. 295–300.
22 P.G. Mezey, Mol. Phys., 1999. **96**(2): p. 169–178.
23 P.G. Mezey, J. Math. Chem., 2002. **30**(3): p. 299–303.
24 W. Kohn and A. Yaniv, Proc. Natl. Acad. Sci. USA, 1978. **75**(11): p. 5270–5272.
25 W. Kohn, Phys. Rev. Lett., 1996. **76**(17): p. 3168–3171.
26 E. Prodan and W. Kohn, *Nearsightedness of Electronic Matter*. 2005: http://arxiv.org/abs/cond-mat/0503124
27 R.F.W. Bader and P. Becker, Chem. Phys. Lett., 1988. **148**(5): p. 452–458.
28 W. Yang, Phys. Rev. Lett., 1991. **66**: p. 1438.
29 W. Yang and T.-S. Lee, J. Chem. Phys., 1995. **103**: p. 5674.
30 C.F. Matta, J. Phys. Chem. A, 2001. **105**(49): p. 11088–11101.
31 P.D. Walker and P.G. Mezey, J. Amer. Chem. Soc., 1994. **116**: p. 12022–12032.
32 P.D. Walker, G.A. Arteca, and P.G. Mezey, J. Comp. Chem., 1991. **12**(2): p. 220–230.
33 N.S. Zefirov and V.A. Palyulin, J. Chem. Inf. Comput. Sci., 2002. **42**: p. 1112–1122.
34 J.S. Murray and P. Politzer, Theor. Comput. Chem., 1998. **5**: p. 198–202.
35 C.M. Breneman and M. Martinov, *The Use of Electrostatic Potential Fields in QSAR and QSPR*, in *Molecular*

Electrostatic Potential: Concept and Applications, J.S. Murray and K. Sen, Editors. 1996, Elsevier: Amsterdam. p. 143–179.

36 J.S. Murray, N. Sukumar, S. Ranganathan, and P. Politzer, Int. J. Quantum Chem., 1990. **38**: p. 611–629.

37 P. Politzer and D.G. Truhlar, *Chemical Applications of Atomic and Molecular Electrostatic Potential*. 1981, New York: Plenum Press.

38 P. Politzer, N. Sukumar, K. Jayasuriya, and S. Ranganathan, J. Amer. Chem. Soc., 1988. **110**: p. 3425–3430.

39 R.G. Parr and W. Yang, J. Amer. Chem. Soc., 1984. **106**: p. 4049–4050.

40 R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. 1989, New York: Oxford University Press.

41 K. Fukui, T. Yonezawa, and C. Nagata, J. Chem. Phys., 1957. **27**(6): p. 1247.

42 K. Fukui, *Theory of Orientation and Stereoselection*. 1975, Berlin: Springer.

43 M.J. Embrechts, J. Robert Kewley, and C. Breneman. *Computationally Intelligent Data Mining for the Automated Design and Discovery of Novel Pharmaceuticals*, in *Smart Engineering Systems: Neural Networks, Fuzzy Logic, Evolutionary Programming, Data Mining and Rough Sets*. 1998. St. Louis, MO: ASME Press.

44 M.J. Embrechts, F. Arciniegas, M. Ozdemir, and M. Momma. *Scientific Data Mining with StripMiner*, in *Proceedings, 2001 SMCia Mountain Workshop on Soft Computing in Industrial Applications*. Blacksburg, Virginia: IEEE Press.

45 R. Kewley, M.J. Embrechts, and C.M. Breneman, *Neural Network Analysis for Data Strip Mining Problems*, in *Intelligent Engineering Systems through Artificial Neural Networks*, C. Dagli, Editor. 1998, ASME Press. p. 391–396.

46 K. Bennett and C. Campbell, SIGKDD Explorations, 2000. **2**(2): p. 1–13.

47 K. Bennett, A. Demiriz, and M. Embrechts. *Semi-Supervised Clustering Using Genetic Algorithms*, in *Artificial Neural Networks in Engineering (ANNIE'99)*. 1999.

48 M.J. Embrechts, F.A. Arciniegas, M. Ozdemir, C.M. Breneman, K.P. Bennett, and L. Lockwood. *Bagging Neural Network Sensitivity Analysis for Feature Reduction in QSAR Problems*, in *2001 INNS – IEEE International Joint Conference on Neural Networks*. 2001. Washington D.C: IEEE Press.

49 R.F.W. Bader, Phys. Rev. B, 1994. **49**(19): p. 13348–13356.

50 R.F.W. Bader and P.M. Beddall, J. Chem. Phys., 1972. **56**(7): p. 3320–3329.

51 R.F.W. Bader, P.M. Beddall, and J. Peslak, Jr., J. Chem. Phys., 1973. **58**(2): p. 557–566.

52 R.F.W. Bader and P.M. Beddall, J. Amer. Chem. Soc., 1973. **95**(2): p. 305–315.

53 R.F.W. Bader and R.R. Messer, Can. J. Chem., 1974. **52**(12): p. 2268–2282.

54 R.F.W. Bader, Acc. Chem. Res., 1975. **8**(1): p. 34–40.

55 R.F.W. Bader, Acc. Chem. Res., 1985. **18**(1): p. 9–15.

56 R.F.W. Bader, M.T. Carroll, J.R. Cheeseman, and C. Chang, J. Amer. Chem. Soc., 1987. **109**(26): p. 7968–7979.

57 R.F.W. Bader, A. Larouche, C. Gatti, M.T. Carroll, P.J. Macdougall, and K.B. Wiberg, J. Chem. Phys., 1987. **87**(2): p. 1142–1152.

58 R.F.W. Bader, J. Chem. Phys., 1989. **91**(11): p. 6989–7001.

59 R.F.W. Bader and C. Chang, J. Phys. Chem., 1989. **93**(13): p. 5095–5107.

60 R.F.W. Bader and D.A. Legare, Can. J. Chem., 1992. **70**(2): p. 657–676.

61 R.F.W. Bader and T.A. Keith, J. Chem. Phys., 1993. **99**(5): p. 3683–3693.

62 R.F.W. Bader, P.L.A. Popelier, and T.A. Keith, Angew Chem. Int. Ed. Engl, 1994. **33**: p. 620–631.

63 R.F.W. Bader, Phys. Rev. B: Condens. Matter, 1994. **49**(19): p. 13348–13356.

64 P.L.A. Popelier, L. Joubert, and D.S. Kosov, J. Phys. Chem. A, 2001. **105**(35): p. 8254–8261.

65 C.E. Whitehead, C.M. Breneman, N. Sukumar, and M.D. Ryan, J. Comp. Chem., 2003. **24**: p. 512–529.

66 C. Whitehead, *Modeling Intermolecular Interaction using the Transferable Atom Equivalent Method*. Ph.D. Thesis, Rensselaer Polytechnic Institute: Troy, NY, 1999.

67 C.M. Breneman and K.B. Wiberg, J. Comp. Chem., 1990. **11**(3): p. 361–373.

68 C.M. Breneman and L.W. Weber, *Transferable Atom Equivalents. Assembling Accurate Electrostatic Potential Fields for Large Molecules from Ab-Initio and PROAIMS Results on Model Systems*, in *The Application of Charge Density Research to Chemistry and Drug Design*, G.A. Jeffrey and J.F. Piniella, Editors. 1991, Plenum Press.

69 C.M. Breneman and T. Thompson, *Modeling the Hydrogen Bond with Transferable Atom Equivalents*, in *Modeling the Hydrogen Bond*, D. Smith, Editor. 1993, ACS Symposium Series: Washington, D.C. p. 152–174.

70 C.M. Breneman, T.R. Thompson, M. Rhem, and M. Dung, Computers & Chemistry, 1995. **19**(3): p. 161.

71 L.H. Hall and L.B. Keir, J. Chem. Inf. Comput. Sci., 1995. **35**: p. 1039.

72 D.E. Clark, J. Pharm. Sci., 1999. **88**(8): p. 815–821.

73 N. Sukumar and C. Breneman, *RECON*. 2001–2006: http://www.drugmining.com/

74 M.J. Frisch, G.W., et al. *Gaussian 98*. 1998, Gaussian, Inc.: Pittsburgh.

75 R.F.W. Bader, *AIMPAC*: http://www.chemistry.mcmaster.ca/aimpac/

76 B.M. Deb, R. Singh, and N. Sukumar, J. Molec. Struct. (Theochem), 1992. **259**: p. 121–139.

77 N. Sukumar, C.M. Breneman, and W.P. Katt. *Virtual high-throughput screening of large datasets using TAE/RECON descriptors*, in *221st National Meeting ACS*. 2001. San Diego: American Chemical Society.

78 M. Randic, J. Amer. Chem. Soc., 1975. **97**: p. 6609–6615.

79 C.M. Breneman, C.M. Sundling, N. Sukumar, L. Shen, W.P. Katt, and M.J. Embrechts, J. Comput. aided Mol. Design, 2003. **17**: p. 231–240.

80 R.J. Zauhar and W.J. Welsh, *Application of the "shape signatures" approach to ligand- and receptor-based drug design*, in *ACS National Meeting*. 2000. Washington, D.C.: American Chemical Society.

81 K. Nagarajan, R.J. Zauhar, and W.J. Welsh, J. Chem. Inf. Comput. Sci., 2005. **45**: p. 49–57.

82 L.H. Hall and L.B. Kier, Bull. Environ. Contam. Toxicol., 1984. **32**: p. 354.

83 L.H. Hall, *Computational Aspects of Molecular Connectivity and its Role in Structure–Activity Modeling*, in *Computational Chemical Graph Theory*, D.H. Rouvray, Editor. 1989, Nova.

84 L.H. Hall and L.B. Kier, eds. *The Molecular Connectivity Chi Indexes and Kappa Shape Indexes in Structure–Property Modeling*. Reviews in computational chemistry II, ed. K.B. Lipkowitz and D.B. Boyd. Vol. 1991. 1991, VCH Publishers. 367–422.

85 L.B. Kier, W.J. Murray, M. Randic, and L.H. Hall, J. Pharm. Sci., 1975. **64**: p. 1971–1974.

86 L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*. 1976, New York: Academic Press.

87 L.B. Kier and L.H. Hall, Eur. J. Med. Chem., 1977. **12**: p. 307.

88 L.B. Kier and L.H. Hall, J. Pharm. Sci., 1981. **70**: p. 583.

89 L.B. Kier and L.H. Hall, *Molecular Connectivity in Structure–Activity Analysis*. 1986, New York: John Wiley.

90 M. Randic, J. Mol. Graph. Model., 2001. **20**(1): p. 19–35.

91 M. Randic, J. Chem. Inf. Comput. Sci., 2004. **44**: p. 373–377.

92 P. Politzer, J.S. Murray, and Z. Peralta-Inga, Int. J. Quantum Chem., 2001. **85**(6): p. 676–684.

93 R.F.W. Bader, P.J. Macdougall, and C.D.H. Lau, J. Amer. Chem. Soc., 1984. **106**(6): p. 1594–605.

94 C.M. Breneman, unpublished.

95 R.F.W. Bader and P.J. Macdougall, J. Amer. Chem. Soc., 1985. **107**(24): p. 6788–6795.

96 P. Sjoberg, J.S. Murray, T. Brinck, and P. Politzer, Can. J. Chem., 1990. **68**(8): p. 1440–1443.

97 C.M. Sundling, N. Sukumar, H. Zhang, M.J. Embrechts, and C.M. Breneman, *Wavelets in Chemistry and Cheminformatics*, in *Reviews in Computational Chemistry*, K.B. Lipkowitz, T.R. Cundari, and V.J. Gillet, Editors. 2006, John Wiley. p. 295–330.

98 C.M. Breneman, N. Sukumar, K.P. Bennett, M.J. Embrechts, M. Sundling, and L. Lockwood. *Wavelet Representations of Molecular Electronic Properties: Applications in ADME, QSPR and QSAR*, in *ACS QSAR in Cells symposium*. 2000. Washington, D.C.: American Chemical Society.

99 M. Wagener, J. Sadowski, and J. Gasteiger, J. Amer. Chem. Soc., 1995. **117**(29): p. 7769–7775.

100 C.M.L. Sundling and C.M. Breneman. *PEST vs. CoMFA: A comparative study of two 3D-QSAR technologies*, in *229th National Meeting, ACS*. 2005. San Diego: American Chemical Society.

101 R. Carbo, J. Arnau, and L. Leyda, Int. J. Quantum Chem., 1980. **17**: p. 1185–1189.

102 R. Carbo and B. Calabuig, J. Mol. Struct. (THEOCHEM), 1992. **254**: p. 517–531.

103 R. Carbo, B. Calabuig, L. Vera, and E. Besalu, Adv. Quantum Chem., 1994. **25**: p. 253–313.

104 R. Carbo, E. Besalu, L. Amat, and X. Fradera, J. Math. Chem., 1995. **18**: p. 237–246.

105 R. Carbo, *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches*. 1995, Amsterdam: Kluwer Academic.

106 R. Carbo-Dorca and P.G. Mezey, *Advances in Molecular Similarity*. Vol. 1, 2. 1996, 1998, Greenwich, CT: JAI Press.

107 R. Carbo-Dorca, D. Robert, L. Amat, X. Girones, and E. Besalu, *Molecular Quantum Similarity in QSAR and Drug Design*. 2000, Berlin: Springer.

108 L. Amat, R. Carbo-Dorca, and R. Ponec, J. Comp. Chem., 1998. **19**: p. 1575–1583.

109 R. Ponec, L. Amat, and R. Carbo-Dorca, J. Comput.-Aided Mol. Des., 1999. **13**: p. 259–279.

110 R. Ponec, L. Amat, and R. Carbo-Dorca, J. Phys. Org. Chem., 1999. **12**: p. 447–454.

111 P.L.A. Popelier, U.A. Chaudry, and P.J. Smith, Perkin Trans., 2002. **II**: p. 1231–1237.

112 P.L.A. Popelier, J. Phys. Chem., 1999. **103**: p. 2883–2890.

113 S.E. O'brien and P.L.A. Popelier, Can. J. Chem., 1999. **77**: p. 28–36.

114 S.E. O'brien and P.L.A. Popelier, J. Chem. Inf. Comput. Sci., 2001. **41**: p. 764–775.

115 S.E. O'brien and P.L.A. Popelier, Perkin Trans., 2002. **II**: p. 478–483.

116 Y.C. Martin, J.L. Kofron, and L.M. Traphagen, J. Med. Chem., 2002. **45**: p. 4350–4358.

117 N. Nikolova and J. Jaworska, QSAR Comb. Sci., 2003. **22**: p. 1006–1026.

118 L. Leherte, N. Meurice, and D.P. Vercauteren, J. Chem. Inf. Comput. Sci., 2000. **40**: p. 816–832.

119 X. Girones, A. Gallegos, and R. Carbo-Dorca, J. Chem. Inf. Comput. Sci., 2000. **40**: p. 1400–1407.

120 X. Girones and R. Ponec, J. Chem. Inf. Model., 2006. **46**(3): p. 1388–1393.

121 N. Sukumar, C.M. Breneman, S.A. Cramer, K.P. Bennett, C.M. Sundling, Q. Luo, and D. Zhuang. *Intelligent Data Mining for Modeling and Prediction of Protein–Ligand, Protein–Surface and Protein–DNA Interactions*, in *Pacifichem 2005: International Congress of Pacific-Basin Societies*. 2005. Honolulu, Hawaii: American Chemical Society.

122 S. Oloff, S. Zhang, N. Sukumar, C. Breneman, and A. Tropsha, J. Chem. Inf. Model., 2006. **46**(2): p. 844–851.

123 M.J. Embrechts, *MetaPLS (Analyze), Version 3.0*. 2001, Rensselaer Polytechnic Institute: Troy, NY.

124 D.V. Eldred, C.L. Weikel, P.C. Jurs, and K.L.E. Kaiser, Chemical Research in Toxicology, 1999. **12**(7): p. 670–678.

125 C. Nantasenamat, T. Naenna, C.I.N. Ayudhya, and V. Prachasittikul, J.

Comput.-Aided Mol. Des., 2005. **19**: p. 509–524.

126 Q. Luo, *Design, Development and Utilization of Novel Molecular Descriptors and Machine Learning Methods*. Ph.D. Thesis, Rensselaer Polytechnic Institute: Troy, NY, 2006.

127 J. Bicerano, *Prediction of Polymer Properties*. Plastics Engineering Series, ed. D.E. Hudgin. Vol. 27. 1996, New York: Marcel Dekker, Inc.

128 C.B. Mazza, N. Sukumar, C.M. Breneman, and S.M. Cramer, Anal. Chem., 2001. **73**: p. 5457–5461.

129 M. Song, C.M. Breneman, J. Bi, N. Sukumar, K.P. Bennett, S. Cramer, and N. Tugcu, J. Chem. Inf. Comput. Sci., 2002. **42**: p. 1347–1357.

130 N. Tugcu, M.H. Song, C.M. Breneman, N. Sukumar, K.P. Bennett, and S.M. Cramer, Anal. Chem., 2003. **75**: p. 3563–3572.

131 C.M. Breneman, C.M. Sundling, N. Sukumar, K.P. Bennett, M.J. Embrechts, and S.M. Cramer. *Beyond PEST descriptors: Binding site and ligand shape/property fingerprints*, in *231st American Chemical Society National Meeting*. 2006. Atlanta, GA: American Chemical Society.

132 C.M. Breneman and N. Sukumar, *New Developments in Molecular Modeling*, in *Yearbook of Science & Technology*. 2004, McGraw–Hill: New York. p. 208–211.

133 C.M. Breneman, K.P. Bennett, M. Embrechts, S. Cramer, M. Song, J. Bi, and N. Sukumar, *Descriptor Generation, Selection and Model Building in Quantitative Structure–Property Analysis*, in *Experimental Design for Combinatorial and High Throughput Materials Development*, J.N. Cawse, Editor. 2002, John Wiley: New York.

19
Fleshing-out Pharmacophores with Volume Rendering of the Laplacian of the Charge Density and Hyperwall Visualization Technology

Preston J. MacDougall and Christopher E. Henze

19.1
Introduction

"When the distribution of charge over an atom is the same in two different molecules, i.e. when the atom or some functional grouping of atoms is the same in the real space of two systems, then it makes the same contribution to the total energy and other properties in both systems." This quotation, from page 3 of Richard Bader's classic monograph [1], displays at once his style and precision of writing, and his desire to get right to the crux of the matter at hand. In this chapter, we examine the charge densities of sets of functional groupings of atoms in drug molecules. These sets are referred to as the "pharmacophores" presumed to be responsible for the pharmacological activity of the drugs [2].

When two things *are* the same, it does not matter how you look at them. As long as you look at them *in the same way*, they will look the same. If two things are *not* the same, particularly if they are different in a subtle way, then their appearance can be quite dependent on how you look at them. For instance, the silhouette of one's right hand, palm down, is nearly indistinguishable from the silhouette of one's left hand, palm up.

An extremely coarse description of pharmacophores is required for rapid screening of a large number of possible drug molecules for variously defined measures of "fitness" with regard to their interaction with a known active site on an enzyme, or some other biological target. Typically, formless spheres are used to represent either inclusion or exclusion volumes, where hydrophobic groups should, or should not be, respectively. Similar, but alternatively labeled, spheres are also used to position hydrophilic groups. Other groups that may be included in a pharmacophore are, but are not limited to:

- positive or negative formal, or partial, charges;
- aromatic rings;
- hydrogen-bond acceptors; and
- hydrogen-bond donors.

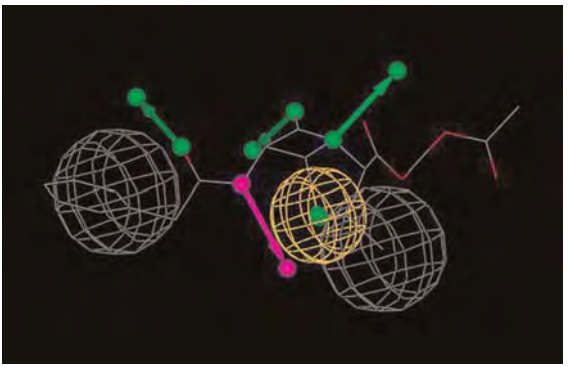


Fig. 19.1 A pharmacophore model for penamecillin. Only those features near the reactive site are included in the model. The green arrows indicate hydrogen bond-acceptor features and the lone mauve arrow indicates the hydrogen bond-donor feature. The gray spheres are hydrophobic exclusion volumes, and the yellow sphere marks the sulfur atom. Created by ''Discovery Studio Visualizer'', Accelrys Software, 2005.

For the last three there may be directional constraints in addition to positional constraints. This reflects the empirical observation that the most stable π–π interactions are face-on, and X–H···Y hydrogen bonds are most frequently linear, or very close to being linear. An illustrative example of a pharmacophore is given in Fig. 19.1, in this case matched by penamecillin, a penicillin derivative.

To draw an anatomical analogy to a pharmacophore, a zoologist does not need to recover all of the bones of an animal to identify the species, its gender, and approximate age, even without doing any genetic testing. Just a finite number of key skeletal components will suffice. Just as there are millions of 75-year-old men but only one R.F.W. Bader (who hopefully still has a long life ahead of him), much more detail is needed in the pharmacophore to winnow from the billions of drug-like molecules those that will have a beneficial effect (let alone the single most effective drug possible).

A common approach in drug design is to start with a ''hit or miss'' approach with regard to the pharmacophore – does a molecule have all the required features within an acceptable distance or not? For most binding sites there will still be a very large number of hits, and from that point different researchers may use different, but almost always proprietary, algorithms to calculate the overall binding energy. There may be an optimum range for this result, and other factors must also be considered, for example solubility in water and fat, conformational flexibility, and permeability through barriers in vivo. Irrespective of which binding algorithm is used, the pharmacophore is obsolete at this point. The entire molecule is fed into a fitting algorithm, all of which are highly approximate by necessity, incapable of *accurately* describing the weak van der Waals interactions that are key to biomolecular interactions.

By ''fleshing-out'' pharmacophores with their highly accurate charge-density distributions, we hope to provide insight into which secondary factors, for example ring size or substituent electronegativity, are most likely to impart subtle but important differences in the reactivities of key functional groups. The long-term plan is that more detail could be added to pharmacophores, but that these would be minimal, and would be informed by model-independent, physical characteristics of the functional groups that they are meant to represent. These refined pharmacophores would yield far fewer ''hits'' on initial screening, potentially enabling high-fidelity quantum mechanical modeling of the binding of all ''first cut'' molecules to the targeted active site, thus yielding a ''final cut'' of much higher quality.

To flesh-out the reactive sites of pharmacophores we have used volume rendering of the Laplacian of the charge density. We previously demonstrated that this graphical technique is very effective for identifying physical features associated with hydrogen-bond-donor sites of different strength, and plainly apparent discrimination between hydrophilic and hydrophobic regions, all without the benefit of ''rules'' [3]. To best identify subtle *differences* between corresponding functional groups of similar, but not identical, molecules we have explored the use of the recently developed hyperwall at NASA Ames [4].

These visualization technologies, and several insights into how pharmacophores might be most efficiently augmented, are discussed in greater detail below.

19.2 Computational and Visualization Methods

19.2.1 Computational Details

All the electron densities discussed here are obtained from ab initio electronic-structure calculations. With the exception of cisplatin, they are all at the Hartree–Fock level, with all-electron basis sets of high-quality (double-zeta plus polarization, or better). The cisplatin charge density was derived from an MP2 calculation employing an effective core potential [5]. All computational details, for example software packages used, basis sets, geometric data, and any optimization constraints, can be found in the references cited.

19.2.2 Volume Rendering of the Laplacian of the Charge Density

It has been amply demonstrated, for large and small molecules and for crystals, and with representation from different regions of the periodic table, that sites of chemical reactivity can be related to the subtle and subatomic ''lumps and holes''

in matter manifested as topological features in the Laplacian of the charge density [6–10].

A wry observation by G.N. Lewis, on receipt of the Franklin Medal, bears repeating here. He wrote "It is always of interest to find that some of our most modern scientific ideas have been vaguely anticipated by scientists of earlier centuries. One of the ideas of Lemery, a contemporary of Robert Boyle, is amusingly discussed in a well known history of chemistry, as follows: 'Yet one of his theoretical conceptions was very odd, and shows how far astray a capable man may wonder, when he deserts observed facts for philosophical speculations. He thought that chemical combination between two substances, such as an acid and a base, might be accounted for by supposing that the particles of the one were sharp, and those of the other porous, and that chemical combination was effected by the fitting of the points into the holes!'" [11].

"Fitting" incurs the need to measure both size and shape. This is easy for a tailor, but topological features in the Laplacian are challenges to the design chemist. It is quite tedious, and of uncertain significance, to exhaustively partition the entire Laplacian distribution into attractor basins, in a manner analogous to atomic partitioning [3, 12]. Qualitative assessments of feature size and shape are easily performed, and are informative, especially within a series of related molecules. These typically employ contour diagrams of the Laplacian, or iso-value surfaces, for example the outer $\nabla^2\rho = 0$ "envelope". Contour diagrams can characterize both size and shape of features, but one "slice" at a time. For molecules of low symmetry, this is troublesome. The utility of Laplacian envelopes is not hindered by low symmetry, but the extent of size and shape characterization is limited by selection of a single value of the Laplacian.

Volume rendering [13] enables a continuous range of Laplacian values, or any other scalar data, to be visualized over a three-dimensional volume of any size. The complete occlusion of *valence shell charge concentration* (VSCC) features "inside" the *valence shell charge depletion* (VSCD) that surrounds all molecules, for instance, is prevented by a tunable "opacity transfer function." In our implementation of this graphical technique [3], this function can be adjusted interactively with superimposed Gaussian and step functions. It has the effect of rendering volume elements in 3D texture memory as opaque, invisible, or with adjustable levels of translucence. The rendering color is keyed to the value of the Laplacian at that point, with warmer colors indicating greater local concentration of electronic charge (white > red > orange > yellow), and cooler colors indicating greater local depletion of electronic charge (violet > blue > green). In much the same way as dye-stained organelles in a cell come into and out of focus under the microscope, topological features that are associated with reactive sites, for example lone pairs, can be probed for size and shape in three dimensions.

The results are useful and visually satisfying, providing a "holistic" method of molecular visualization that not only resonates with what chemists intuitively imagine molecules to be like, but have appealing aesthetic qualities that garishly-colored mirrored balls and golden rods do not [14]. Figure 19.2 is a workstation screenshot of the purine base adenine being rendered with our volume

rendering software, nicknamed EVolVis, for Electronic Volume Visualizer. Minimum hardware requirements, and details of the program's algorithm, which uses OpenGL and the X Window system, are given in Ref. [3].

Those who are familiar with Laplacian distributions, will immediately see that this one image provides more information than multiple Laplacian envelopes, or a very large number of contour diagrams, all of which must be mentally collated to furnish a full 3D mental image of the molecule. The frame shown yields simultaneous rendering of lone pair features, for example on the three imine nitrogen atoms (bulbous red protrusions), the deep π charge depletions above and below the carbons (blue balls), and the shallow charge depletions behind the amino hydrogen atoms (little green "yarmulkes"). Adjusting the appropriate slider bar can "grow" the orange VSCCs (new growth becoming "cooler" in color) until the π holes "close". The last carbon to do so must have the greatest local depletion at the corresponding ring CPs, thus implying the greatest susceptibility to attack by nucleophiles.

Figure 19.3 is a volume rendering of $\nabla^2\rho$ for penamecillin, the molecule that was shown schematically in Fig. 19.1. The color key is as in Fig. 19.2, but the transfer function has been adjusted to contrast the hydrophilic diester chain (top right) and the hydrophobic aromatic ring (bottom left), and to bring into focus the key features on the perimeter of the reactive pocket (bracketed by the aromatic ring). There are 22 hydrogen atoms in this molecule, but only one is predicted by the "N, O, F rule" to be a hydrogen-bond donor. This amide hydrogen is jutting into the reactive pocket and is, indeed, revealed to have the only small charge depletion (green yarmulke) that is seen at the "appropriate" sites in adenine (Fig. 19.2). The sulfur atom also borders the reactive pocket, and its twin lone pair charge concentrations (orange) are rendered quite differently from those of any of the oxygen atoms. One of the sulfur atom's large charge depletions is also exposed to the reactive pocket, creating a chemically anisotropic environment for whatever fills it. This anisotropy is, literally, a highly varying "texture" which is naturally rendered by the Laplacian of the charge density.

The pharmacophore in Fig. 19.1, along with the rest of the molecule, is fleshed-out in Fig. 19.3, but it is impossible to attribute pharmacological significance to the physical features that are revealed. Comparing the full, fleshy details of the pharmacophore in an active drug with those in an inactive drug would be an excellent opportunity to learn something about the pharmacological significance of the sizes and shapes of topological features in the Laplacian.

Wagner et al. have reported crystallographic studies of penamecillin and its partially oxidized sulfoxide (the only difference is that one of the lone pairs is replaced with an oxo ligand) [15]. Local oxidation can be expected to alter the electrostatic potential in the region of oxidation, which was indeed observed by Wagner et al., in addition to some conformational differences.

We repeated the volume rendering procedure for the second penicillin derivative, again using the nuclear coordinates from the crystallographic investigations. Again, as might be expected, there are large areas of near-perfect transferability far from the site of oxidation, whereas there are gross differences at the

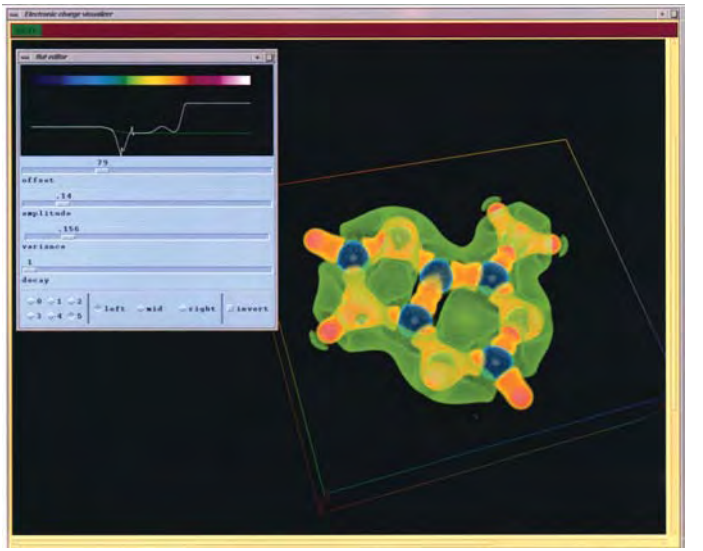


Fig. 19.2 A screenshot of adenine "under the microscope" of volume rendering of the Laplacian of the charge distribution. The white line in the pop-up window is the overall opacity transfer function that produced the image shown. It is composed of simple Gaussian and step functions that are individually tuned with the slider controls shown. Color corresponds directly to the value of the Laplacian, with white being the most negative and violet being the most positive. The same general color scheme is used in all other volume renderings in this work. Adapted from Ref. [3].

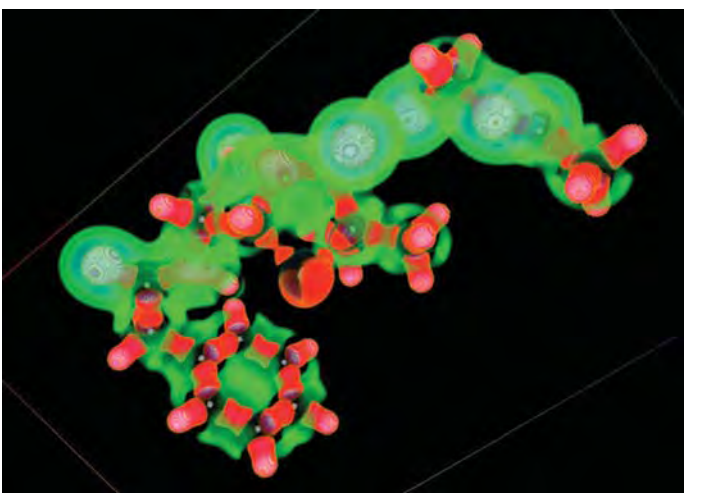


Fig. 19.3 Volume rendering of $\nabla^2\rho$ in penamecillin. Adapted from Ref. [3].

sulfur. These same observations would be made by comparing the minimally-informative pharmacophores that correspond to the two molecules. We are interested in the subtle, but chemically significant details that may be evident in the fleshed-out pharmacophores. By *sequentially* volume rendering the Laplacian distributions of the two molecules, such differences can be identified, but we are back in a similar situation as before – having to mentally contrast separate renderings. Truly parallel renderings would be preferable.

19.2.3 The Hyperwall

The hyperwall [4] is simply a 2D array of flat-screen displays with each display driven independently by a graphics workstation and with all the workstations connected by a network switch. Single large images can be rendered on the aggregate display (''powerwall mode''), but we have found it generally more useful to display sets of related images (''hyperwall mode''). The sets can be laid out according to a single identifier or a pair (row, column), and the set members (slaves) may be functionally connected to each other and/or to a controlling terminal (master node) over the intervening network. This arrangement provides a large number of pixels (65 million in the current implementation) but retains a high level of system support (CPU, GPU, memory, disk) per pixel, so that multiple instances of resource-intensive graphics techniques, for example time-varying volume visualization, can be efficiently deployed in parallel. The simultaneous display of, and interaction with, such visualizations provides a high-bandwidth, information-rich exploratory environment that enhances the remarkable power of our own visual systems for detecting differences, trends, outliers, and subtle patterns.

A powerwall is simply a very large image made up of multiple computer screens – the total information content is equal to the sum of its parts. By relating the information content of the screens, in a user-controlled manner, the hyperwall is a demonstration of the teamwork maxim.

19.2.4 Hyper-interactive Molecular Visualization

In Fig. 19.4, what distinguishes the hyperwall from a powerwall is not apparent. For one thing, it is a static picture. On a powerwall the screens can simply display distinct video streams, but on the hyperwall ''blocks'' of screens, or the entire 7×7 array, can be synchronized. It is also possible for the user's interaction with one screen, or a row/column of screens, to refresh the data displayed on other screen(s). For example, one column could be a series of related molecules, and the cursor on the ''master node'' (left, foreground in Fig. 19.4) could move a small cube in a concerted manner within each molecule. Another column could display 3D scatter plots of data within the cube on the molecule of the same row.

Fig. 19.4 The NASA Ames hyperwall with parallel interactive volume rendering of two derivatives of penicillin – one active, one inactive. One of us (PJM) is clearly excited. The upper row has volume renderings of the Laplacian for penamecillin. The lower row has corresponding images for the inactive sulfoxide derivative. Starting from the left, the columns are: entire molecules, in the conformation observed in the corresponding crystal (note that the reactive pocket is not bracketed by the aromatic ring in the inactive form); β-lactam groups present in all penicillin derivatives (the blue halo is around the carbonyl oxygen and the reddish lumps on the right is the lone pair on the tertiary amino nitrogen); the reactive pocket with amide hydrogen in the middle; sulfur (upper) or sulfoxide group (lower); diester side-chain; aromatic ring (note how nearly perfectly transferable it is). High-resolution images of each of these screens is available online at www.nas.nasa.gov/Groups/VisTech/hyperwall/ Select ''Papers, etc.'' in the left menu, and then ''screendumps2'' in the index. Clicking on the top penicillin thumbnail will bring up a screenshot similar to Fig. 19.4. Then, clicking on any hyperwall screen will link to a large, high-resolution image of that particular screen.

For our purpose of identifying subtle differences in the real functional groups that constitute the pharmacophore models of related drug molecules, we use rows of ''slaves'' for different molecules, and columns of slaves for functional groups, including those in the pharmacophore. The conformational differences in the derivatives is apparent in column 1, In column 6, the two screens appear almost identical, even though the phenyl rings are oriented quite differently. In column 5 the diester side-chains are conformationally different but reactive features are rendered synchronously as the transfer function sweeps different parts of the Laplacian spectrum. The differences between the rendering of the lone pairs on keto and ether oxygen atoms are more pronounced than similar features in different rows.

More contrast is seen in columns 2, 3, and 4 as the rendering searches for reactive features. The ''blocking'' of columns, individually or in groups, enables optimum transfer functions for different elements of the pharmacophore. An mpeg file with a sample video of interactive rendering on the hyperwall, can be downloaded at www.nas.nasa.gov/~chenze/preston/pen2.mpg.

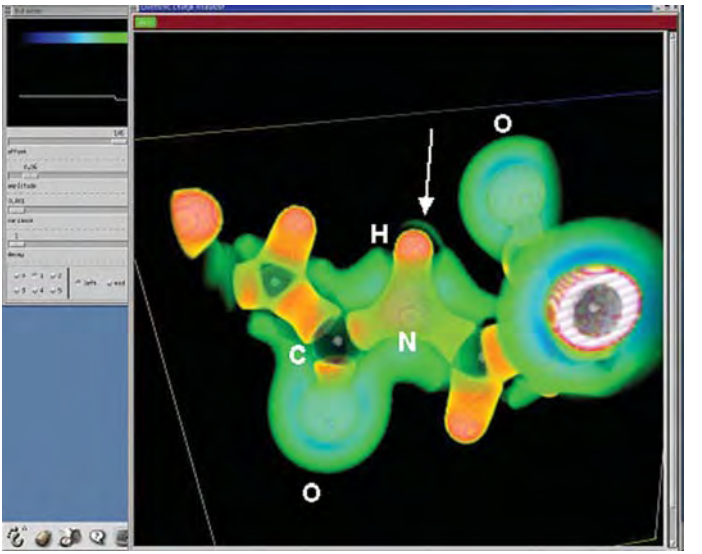


Fig. 19.5 A volume rendering of $\nabla^2\rho$ in the reactive pocket of the inactive penicillin derivative. The local charge depletion behind the amide hydrogen (green yarmulke, indicated by the arrow) is seen pulled to the right, interacting with the nearby oxo group in that direction (some key atoms in this region are identified). This image is a screenshot of column 3, row 2, from Fig. 19.4.

One possible reason for the inactivity of the oxidized derivative is revealed in column 5. The hydrogen-bond donor feature that is *fully exposed* in the reactive pocket of the active form (Fig. 19.3) is severely deformed by the nearby oxygen atom of the added oxo group (Fig. 19.5). Intra-molecular hydrogen bonding thus could interfere with this molecule's pharmacological activity.

Filling the hyperwall with fleshed-out pharmacophores, and scrolling through libraries of drug molecules, all the while searching for insights such as those above, is an exciting prospective use of this combination of visualization techniques.

19.3 Subatomic Pharmacophore Insights

19.3.1 Hydrogen-bonding Donor Sites

In the first systematic study of hydrogen bonding orientations on the basis of topological properties of the Laplacian, Carroll et al. used HF as donor and a variety of bases as acceptors [16]. They found excellent agreement with reported geome-

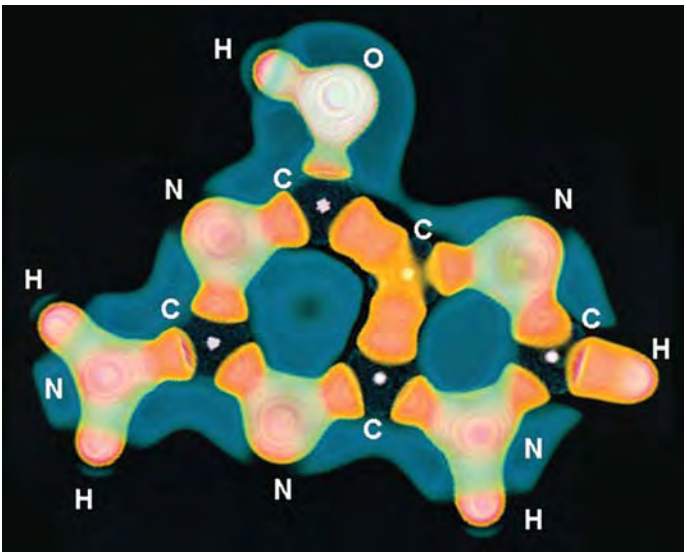


Fig. 19.6 Volume rendering of $\nabla^2\rho$ in 6-hydroxy-2-aminopurine. A different color key was used, but the ordering is the same as that discussed in the text. Adapted from Ref. [3].

tries of such van der Waals complexes when the local charge depletion behind the hydrogen atom of HF was collinearly aligned with the nucleus of the base atom and its nonbonded charge concentration.

We have found that this topological feature of the donor hydrogen atom's VSCD is usually not present when the X–H is not a rule-based donor hydrogen bond, and that, when present, its magnitude parallels that of expected hydrogen-bond-donor strength [3]. All the X–H bonds in Figs 19.2 and 19.3 are consistent in this regard. We note that it is not simply a rendering artifact. The VSCD topology behind the hydrogen atoms of most C–H bonds, and X–H bonds in general when X is less electronegative than H, is *opposite* to that of the yarmulkes. They are local $\nabla^2\rho$ minima *within* the VSCD whereas the yarmulkes are local maxima.

The Laplacian for the hydroxy isomer of guanine is volume rendered in Fig. 19.6. There are four types of X–H bonds, and the presence and size of the yarmulkes behind each is consistent with its expected hydrogen bonding capacity or strength. As expected, the amino and imino features are very similar.

Within QTAIM, identification of reactive sites, and their characterization, is performed by topological analysis of $\nabla^2\rho$ by virtue of its role in describing the energetics of many-electron systems in stationary states [1]. Such a model-independent procedure cannot be expected always to conform to established rules. Figure 19.7 indicates that hydrogen-bond donor features are not only found ''where expected''. ''Weak hydrogen bonds'', that do not follow the standard ''rules'', are believed to frequently play important structural and mechanistic roles in biological processes. Important classes of such hydrogen-bond donors are C–H

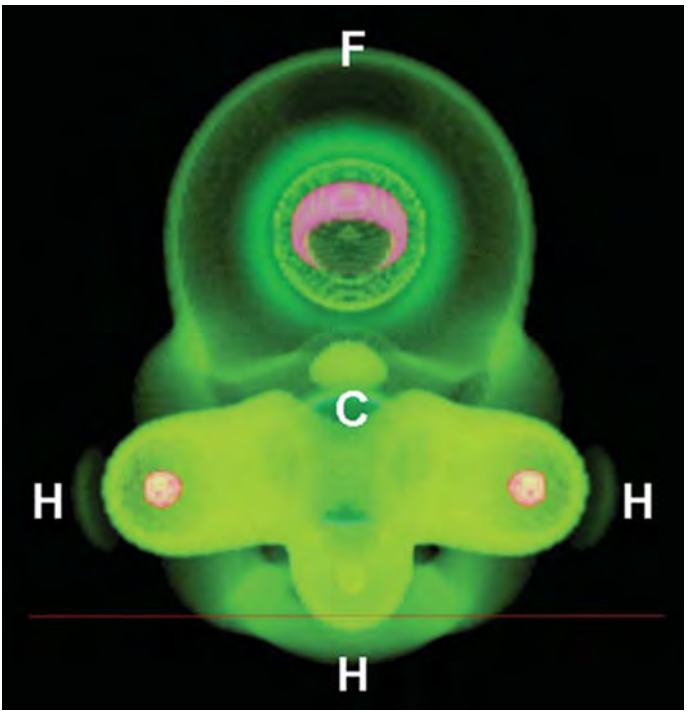


Fig. 19.7 Volume rendering of $\nabla^2\rho$ in methyl fluoride. Note the barely formed yarmulkes beyond each of the ''activated'' hydrogen atoms. The pink halo is the torus of nonbonded charge concentration corresponding to the three lone pairs on fluorine.

bonds in which the carbon atom is ''activated'' by a highly electronegative substituent, for example fluorine, or oxygen in aldehydes. Corresponding to weak H-bond donor strength, the yarmulkes behind the hydrogen atoms in methyl fluoride are barely formed (Fig. 19.7), but they are a topological change, relative to simple alkanes, induced by fluorine substitution.

With systematic studies, the existence, strength, and direction of hydrogen-bond donor features in pharmacophore models could be simply augmented with knowledge gleaned from topological analyses of $\nabla^2\rho$. Similar pharmacophore improvement is envisaged for hydrogen-bond acceptor features, based on analogous topological studies of nonbonded charge concentrations. Such studies have been well-reported in the literature [6–10].

19.3.2 Inner-valence Shell Charge Concentration (i-VSCC) Features in Transition-metal Atoms

Within the context of the orbital model of electronic structure, when explaining the ground states of elements in the d-block of the *n*th row of the periodic table,

we must point out to our students that there are exceptions to the aufbau principle. Depending on electron count, electrons may occupy either (n)s or $(n-1)$d atomic orbitals. There are assorted reasons given for this orbital jockeying, but the similarity of the energy levels of these nonphysical one-electron states always features.

In the minds of students the energy overlap often leads to the incorrect assumption that (n)s and $(n-1)$d orbitals also have comparable radial distributions. In fact, the radial distribution functions of $(n-1)$d orbitals peak at radii much closer to those of an $(n-1)$s orbital in the same atom, which is considered to be part of the "core".

In many-electron reality an electron is just an electron. For an atom in the d-block of the periodic table the local concentration maxima in the outermost shell of charge concentration have significant contributions from what are *ostensibly* core and valence orbitals. Since these concentrations have been shown to have stereochemical consequences and to effect metal–ligand bonding [17–19], the outermost shell of charge concentration in the "core" is called the inner VSCC, or i-VSCC.

Figure 19.8 is a good illustration of a common topological relationship between the i-VSCC of a transition-metal atom and the dative bonding charge concentrations in the ligand sphere – lumps fitting into holes or grooves. The high angular momentum of electrons at the radii of the holes allows them much sharper definition than is seen in the π holes of organic compounds. This imparts stronger

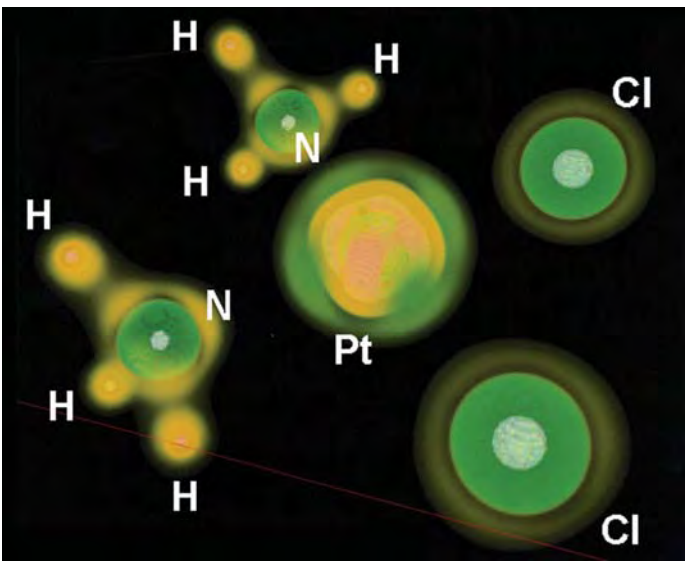


Fig. 19.8 Volume rendering of $\nabla^2\rho$ in cisplatin. A different color key was used, but the ordering is the same as discussed in the text. Adapted from Ref. [3].

directionality to metal–ligand interactions than is seen in groups IA and IIA, without invoking covalence.

Because many enzymes have transition-metal ions at or near their active site, factors affecting optimum coordination directions for different metals, and in their different oxidation states, may aid pharmacophore development. In much the same way as a hydrogen-bond donor feature in a pharmacophore includes a single, collinear arrow (Fig. 19.1), a Pt(II) feature could include four variously directed arrows. Ligand field theory suggests that they should be at 90° angles, anchored by permanent ligands, but volume rendering of $\nabla^2\rho$ then topological analysis could refine these directions. These latter investigations may seem unnecessary if one assumes that the geometries of countless coordination complexes already imply optimum directions of metal–ligand bonding for different metal ions in their different oxidation states. This, however, is not always a valid assumption.

19.3.3
Misdirected Valence in the Ligand Sphere of Transition-metal Complexes

The geometry of chromyl chloride, CrO_2Cl_2, is "opposite", in a VSEPR model sense, from that of sulfonyl chloride, SO_2Cl_2. Both have distorted tetrahedral geometries, but the angle between the double bonds to oxygen are is *smaller* than the tetrahedral angle when the central atom is chromium, whereas it is larger (as expected) when the central atom is sulfur.

This and other geometric conundrums involving transition-metal atoms continue to be clarified by computational and experimental examination of the topological properties of the i-VSCC of the metal atom [17–20]. One reoccurring curiosity is that of misdirected valence.

Liehr coined the term "misdirected valence" in 1964 to improve the ligand theoretical predictions of the spectra of chelation complexes and other metallacycles [21]. If one ignores the metal atom that is chelated by a diamine, for instance, the shape around each nitrogen will be trigonal pyramidal, and the principal axes will be directed in the general direction of the metal atom. The "bite angle" of the ligand, and the size and d-electron count of the metal ion, will determine how "well-aimed" these axes are. Simple geometry-based corrections for misdirected valence greatly improved the accuracy of ligand field theory in such complexes [21].

Figure 19.9 presents stark physical, not inferred, evidence of misdirected valence. In CrO_2Cl_2, the stereochemical effect of the ligand-opposed charge concentrations of chromium's i-VSCC [18, 19] has narrowed the OCrO angle. In Fig. 19.9b the approximate twofold axis that is typically collinear with the M=O bond, and bisects the nonbonded charge concentrations in the VSCC of oxygen, is strongly askew. As indicated in Fig. 19.9b by the double-headed arrow, the local twofold axes, with regard to the structure of the VSCCs of the two oxygen atoms, are not directed at the chromium *nucleus*, but at the single charge depletion in the i-VSCD that apparently serves as the point of attraction for the oxygen ligands [18].

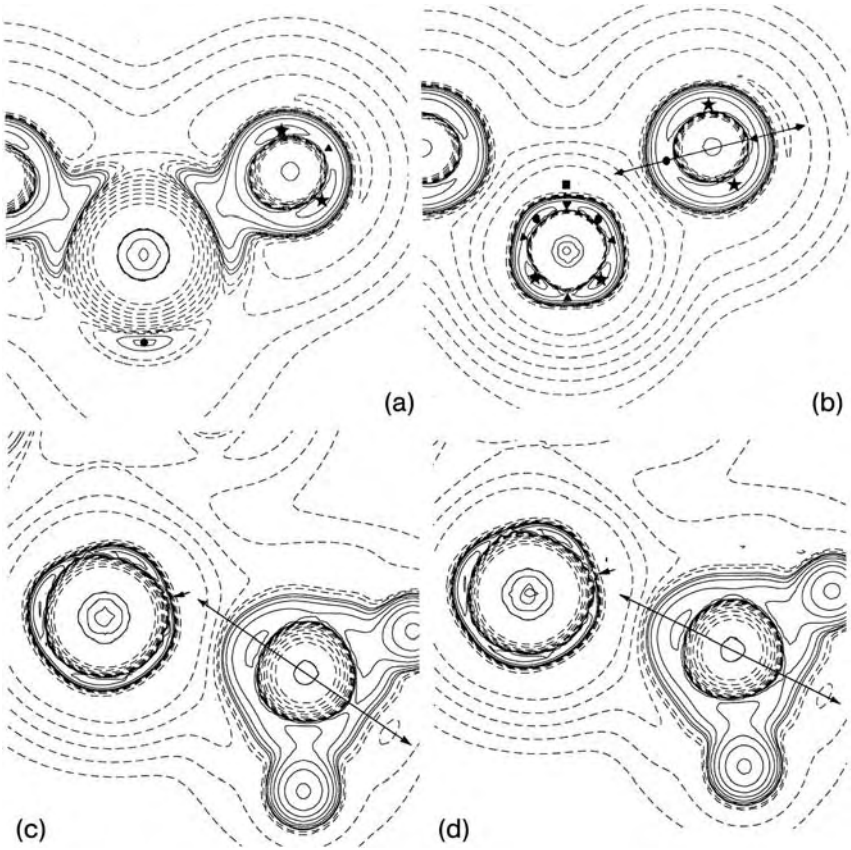


Fig. 19.9 Contour plots of the Laplacian of the charge density for SO_2Cl_2 (a), CrO_2Cl_2 (b), and $H_2CNbCl_2C_5H_5$ (c and d). The solid (dashed) contours indicate negative (positive) Laplacian values, and the contour values are spaced logarithmically (jaggedness is a grid artifact). Plots a and b are centered on the mid-point of the S=O (a) and Cr=O (b) bonds, and the stars (squares) indicate maxima (minima) in $-\nabla^2\rho$. Plots c and d are centered on the mid-point of the Nb=C bond. The computed equilibrium geometry of this molecule is between these two "rocked" geometries. In b, c, and d, the arrows indicate approximate local twofold rotational axes.

Also in Fig. 19.9, the strong asymmetry in the C–R bonds of a Schrock carbene [22] is evidently another example of misdirected valence and not a result of an agostic interaction between a metal ion and a C–H bond. Here the metal is a niobium(III) ion in $H_2C{=}NbCl_2C_5H_5$. The double bond in such Schrock carbenes couples a carbene, for which triplet ground states are favored, to d2 metal ions, for which, again, a triplet state is favored.

As discussed earlier, for the general case of d-block atoms, the two 4d electrons are regarded as being chemically valent but as making their physical contribution primarily to the outermost core shell of charge concentration (in this case the 4th). Also contributing largely to this "outer core" shell of charge concentration are the eight 4s and 4p (truly) core electrons. The topology of the i-VSCC of the

niobium atom has, as one might expect in an extended Lewis model, four large and two smaller concentrations. The local maxima in the Laplacian form a distorted trigonal prism whose four-sided faces are perpendicular to the plane shown in Fig. 19.9 (c and d) [18].

The cyclopentadienide ligand occupies such a large fraction of the ligand sphere that, again, there is a forced misdirected valence. To demonstrate that the ''rocking'' of the carbene is not caused by a M···H–C attraction, we depict the remarkable invariance of *both* the i-VSCC of niobium *and* all the charge concentrations of the carbene, for pivotal 5° deviations about the equilibrium geometry. There is no notable distortion of the niobium i-VSCC, nor ''agostic'' C–H bond-charge concentration until the carbene has been rocked much further along this constrained reaction coordinate [23].

Allowing for such misdirected valence in pharmacophore definition would not be a simple matter. For example, steric factors within the ligand sphere of the metal atom would have to somehow mesh with its optimum coordination directions. The most important modification would be to *shift* the focal point of ligand attraction *from* its assumed position at the metal's nucleus, *to* inner valence shell topological features.

19.4 Conclusion

Rapid screening in the vast virtual space of chemical compounds and all their conformations requires extremely high selectivity to be of any practical use. For quality results the selection tools must also be accurate. There can be no doubt about the fineness and utility of the set of tools for characterizing the structural and atomic properties of matter that have been developed over the decades by the ''Bader school''. These tools have a rigorous theoretical foundation and are backed by experimental results. Richard Bader has seen the future and, in it, computational chemists ''synthesize'' arbitrarily large and complex molecules like nature does it, cutting and splicing together the most convenient fragments, not ''from the beginning'' every time [24].

Matta has demonstrated the feasibility of this approach for known drug molecules [25]. Breneman et al. [26] have developed the transferable atom equivalent (TAE) method that can rapidly predict key properties of large molecules from a finite set of atomic descriptors that have been obtained from QTAIM analysis of smaller molecules. Indicating its potential utility in drug design, this method has had some success in modeling van der Waals interactions [27]. Breneman's TAE approach is reviewed in Chapter 18 of this book.

Until Bader's vision is achieved, drug screening tools can be incrementally improved in a manner that is consistent with the proven success of the Bader school's philosophy – seek advances based on rigorous theory and back them up with experiment. We have demonstrated ways in which modern computer graphics techniques can be used to extract physical and insightful information about pharmacophore features. We have also suggested practical ways in which

pharmacophore features could be augmented with data that can be obtained from computation or experiment. These suggestions may be particularly beneficial because they all relate to noncovalent interactions, which are key to pharmacological activity and are a weakness of current high-throughput ab-initio methods.

References

1 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Clarendon Press, Oxford, 1990.

2 P. Gund in *Pharmacophore Perception, Development, and Use in Drug Design* (Ed.: O. F. Güner), IUL Biotechnology Series, International University Line, La Jolla, CA, 2000, Part I.

3 P. J. MacDougall, C. E. Henze, *Theor. Chem. Acc.*, 2001, *105*, 345–353.

4 T. A. Sandstrom, C. E. Henze, C. Levit, "The hyperwall", in *Coordinated and Multiple Views in Exploratory Visualization*, IEEE Conference Proceedings, 2003, 124–133.

5 P. N. V. Pavankumar, P. Seetharamulu, S. Yao, J. D. Saxe, D. G. Reddy, F. H. Hausheer, *J. Comput. Chem.*, 1999, *20*, 365–382.

6 R. F. W. Bader, P. J. MacDougall, C. D. H. Lau, *J. Amer. Chem. Soc.*, 1984, *106*, 1594–1605.

7 R. F. W. Bader, P. J. MacDougall, *J. Amer. Chem. Soc.*, 1985, *107*, 6788–6795.

8 R. F. W. Bader, *Chem. Rev.*, 1991, *91*, 893–928 (and references therein).

9 P. Coppens, T. Koritsanszky, *Chemical Reviews*, 2001, *101*, 1583–1627 (and references therein).

10 P. L. A. Popelier, P. J. Smith, "Quantum Topological Atoms", in *Chemical Modelling: Applications and Theory, Vol. 2*, Royal Soc. Chem., London, 2002, Chapter 8 (and references therein).

11 G. N. Lewis, "Acids and Bases", *J. Franklin Inst.*, 1938, *226*, 293–313.

12 P. L. A. Popelier, N. O. J. Malcolm, *Faraday Disc.*, 2003, *124*, 353–363.

13 R. A. Drebin, L. Carpenter, P. Hanrahan, "Volume rendering" in *Computer Graphics* (Ed.: J. Dill), SIGGRAPH '88 Proceedings, 1988, *22*, 65–74.

14 M. Kemp, *Nature*, 2001, *412*, 588. (The mpeg file that is discussed in this article, a lovely "pirouette" of vitamin B12, can be downloaded from www.nas.nasa.gov/~chenze/preston/b12.mpg)

15 A. Wagner, R. Flaig, B. Dittrich, H. Schmidt, T. Koritsanszky, P. Luger, *Chem. Eur. J.*, 2004, *10*, 2977–2982.

16 M. T. Carroll, C. Cheng, R. F. W. Bader, *Mol. Phys.*, 1988, *63*, 387–405.

17 P. J. MacDougall, M. B. Hall, R. F. W. Bader, J. R. Cheeseman, *Can. J. Chem.*, 1989, *67*, 1842–1846.

18 P. J. MacDougall, M. B. Hall, *Trans. Amer. Cryst. Assoc.*, 1990, *26*, 105–123. P. J. MacDougall, M. B. Hall, unpublished results, 1991.

19 I. Bytheway, R. J. Gillespie, T.-H. Tang, R. F. W. Bader, *Inorg. Chem.*, 1995, *34*, 2407–2414.

20 W. Scherer, P. Sirsch, D. Shorokov, M. Tafipolsky, G. S. McGrady, E. Gullo, *Chemistry*, 2003, *9*, 6057–6070.

21 A. D. Liehr, *J. Phys. Chem.*, 1964, *68*, 665–722.

22 R. R. Schrock. *Acc. Chem. Res.*, 1979, *12*, 98–104.

23 P. J. MacDougall, M. B. Hall, unpublished results, 1991.

24 C. Chang and R. F. W. Bader, *J. Phys. Chem.*, 1992, *96*, 1654–1662.

25 C. F. Matta, *J. Phys. Chem. A*, 2001, *105*, 11088–11101.

26 C. M. Breneman, T. R. Thompson, M. Rhem, M. Dung, *Computers and Chemistry*, 1995, *19*, 161–179.

27 C. M. Breneman, M. Rhem, *J. Comp. Chem.*, 1997, *18*, 182–197.

Index

a
ab initio wavefunction 86
ab initio periodic approaches 169
Abramov functional 432
Abramov energy density 432
absolute hardness 400
absorption intensity 79
accuracy of atomic integrations 17, 18
acenaphthylene 408
acetamide 333
acrolein 129, 135, 136
action integral 41
action principle 39, 40, 44
active site 511
additivity of the atomic energy 10
adenine 312
adenine binding site 305 ff., 308, 312, 313
adenine binding site of the protein hAR 302, 303, 310
adenosine 323
ADI 405, 406
agostic interaction 364, 366, 367, 512
agostic systems 365
AIM2000 27, 101, 103 ff.
AIMALL97 28
AIMDELOC 15, 28
AIMPAC 27, 29, 264
Al 216
$Al_2C_2H_{10}$ 145
Al–Al bond 211
aliphatic hydrocarbons 379
alkali halides 214
alkali metal clusters 192
alkaline-earth metals 214
alkane 77, 91, 109, 116 f.
n-alkanes 99
n-alkenes 194
alkyne 114
amino acid 21, 22, 24, 301, 324, 325
amino acid residues 56, 285, 302, 303, 327
ammonia 276
analytical CPHF method 101
anharmonicity 98
anomeric effect 375, 386
anthracene 408
aromatic compounds 399
aromatic radicals 393
aromatic stabilization energy 400
aromatic transition state (TS) 418
aromatic π-sextets 416
aromaticity 394, 395, 399
– indices 394, 418 f.
asphericity shifts 334
atom types in proteins 26
atomic additivity schemes 82
atomic basin 6, 211, 428
– shape 209 ff.
– volume 16, 295
atomic charge 9, 16, 75, 328, 401
– derivatives 80, 81
– transfer derivatives 80
– transfer dipole contribution 75
atomic components 176
atomic continuity theorem 51
atomic contributions 50, 62, 65, 71, 73, 74, 76, 81, 90 ff., 108
– to $\bar{\alpha}$ 112
– to $\Delta\bar{\alpha}/\Delta r_{CH}$ 112
– to the electronic energy 18ff., 276
– to the magnetizability tensors 91
– to the polarizability tensors 77
atomic current theorem 51, 52
atomic dipolar polarization 20
– contribution 72
– gradient 75
atomic dipoles 109
atomic electric polarizability tensors 91

atomic electron populations 16, 103, 175, 277, 401
atomic electronic energy 19, 84
atomic electronic kinetic energy 17, 82
atomic electronic virial theorem 19, 83
atomic electrostatic multipole moments 20–25, 126
atomic energy XXII, 20, 29, 84
atomic exchange matrices 138
atomic force
– microscope 54
– theorem 51
atomic fragments 317
atomic graph 348
atomic group source function 195
atomic interaction lines 8
atomic interactions 260, 347
atomic magnetic polarization contribution 88
atomic multipolar ED 296
atomic multipole moments 20–25, 126
atomic multipoles model 122
atomic net charges 16, 188, 295, 299
atomic net current contribution 88
atomic nuclear virial energy 82 ff., 85
atomic polarizability 77
atomic polarization 20–25, 78, 80
atomic populations *see* atomic electron population
atomic power theorem 51
atomic properties 15–25, 26, 194
atomic quadrupolar polarization 24
atomic quadrupole moment 24, 175, 352
atomic source contributions in $Mn_2(CO)_{10}$ 200
atomic surface derivative contributions 76
atomic theorems 51
atomic torque theorem 51
atomic virial theorem XXII, 19, 51
atomic volume 16, 277, 328, 352
attraction (repulsion) basin 210
Au (111) surface 240
AX_4 (CH_4, CF_4, $SiCl_4$) 154
AX_6E molecules 155 f.
AX_7 molecules 155 f.

b

B_2H_6 145
back bonds 180
back-bonding 363
back-donation 363, 364
π-back-donation 352
basin populations 16, 154
basin volume 16, 297, 299
basins 142
basis set superposition error (BSSE) 460, 461
BCC *see* bonded charge concentration
BCP *see* bond critical point
Becke and Edgecombe 438
benzene 392, 393, 395, 403, 406, 408, 412
benzocyclobutadiene 408
m-benzyne 392 ff.
BF_3 6 ff.
BH_3 368
bicyclo-[1.1.1]-pentane 98, 100, 101, 111, 116
bicyclo[3.3.1]nonane 115
bicycloalkane 99, 114, 116
binding energy 454
biphylene 408
bond charges 72, 75, 80, 275
bond contribution 66 ff., 69 ff., 74, 92
– to the energy-gradient-based force 85
bond critical point (BCP) 4, 10, 25, 141, 168, 231, 260, 271, 375, 426, 431 ff., 477
bond current contribution 89
bond ellipticity 12, 184, 185
bond force 85
bond order 11, 13, 15, 25, 135
bond path 4, 11, 52, 141, 232, 260, 386
– connectivity 357
bond polarizability model 97
bond properties 11, 218
bonded charge concentration (BCCs) 182
bonded interactions 130
bonded radius 11
bond-length aberrations 334
bond-order model 425
borazine 407
Born–Oppenheimer energy surface 460
Born–Oppenheimer approximation 123
Born–Oppenheimer procedure 55
Bragg angle 285
branched alkane 114
BSSE *see* basis set superposition error
Buckingham-type potentials 437
buckminsterfullerene allotropes of carbon 208
buckybowls 409
bulk modulus 220, 221
1,3-butadiene 418
butane 111, 115, 129, 194
2-butanol 332

c

C_2H_2 156
C_2H_2 dimer 129

$(C_2H_2)_2$ 132
C_2H_4 156
C_2H_6 348
CaF_2 216
cage critical point 4, 5, 168, 388
calculus of variations 41
Carbo index of similarity 26, 487
carbon monoxide 20, 21, 348, 445
carbonyl supported metal–metal interactions 357
Cash–Karp Runge–Kutta (CRRK) method 232
catalytic activity 236
catastrophe point 357
CH bond 111
CH stretching modes 96
CH_4 103, 109
characteristic set 4
charge-assisted hydrogen bonds 199, 458
charge concentrations 141
charge-density refinement 308
charge transfer 78, 176, 188, 431
– contribution 73, 80
– dipole 21
– moments 176
chemical structure 52
chemical transferability 195
chiral invarioms 330
chiral modules 336
chirality 336
chlorine 276
chrysene 408
ClF_3 155
cisplatin 510
Clar structure 417
Clar's aromatic sextet 395, 416f.
class I clathrates 189
clathrate type I structure 187
closed-shell bonding 8, 12
closed-shell character 355, 445
closed-shell interactions 11, 174, 260, 358, 427, 436, 441, 443
clustering procedure 313
CO *see* carbon monoxide
$CO(NH_2)_2$ 264, 272
CO_2 82
$Co_2(CO)_6(ASH_3)_2$ 352, 353
$Co_2(CO)_6(NH_3)_2$ 353
$Co_2(CO)_6(PH_3)_2$ 353
$Co_2(CO)_6(XH_3)_2$ 352
$Co_2(CO)_6(\mu\text{-}CO)(\mu\text{-}C_4H_2O_2)$ 355
$Co_2(CO)_8$ 201, 369
$[(CO)_5Cr{-}H{-}X]^-$ 368
$(CO)_5M(H_2)$ 364
commutator average 47
compensatory transferability 56
complementary ligands 488
π-complexes 357 ff.
π-complex–metallcycle 364
compressibility 220, 221
conditional pair density 10, 192
conformational space 357
connectivities 207
contergan/thalidomide scandal 336
CONTOR 29
coordination 351
coordinatively unsaturated sites (CUS) 236
core basin 142
core-valence bifurcation index 439
correlation energy 276, 277
correspondence rule 345
Coulomb
– correlation 14, 377
– energy 123, 125, 134, 139
– expansion 131
– hole density 402
– interaction 125, 129, 133
coupling constants 375
covalence degree (DC) 447
covalent bonds 216, 354, 455
CPHF method 80, 91, 106
CPU time 129
$Cr(CO)_5$ 368
$Cr(CO)_5(C_2H_4)$ 363
$Cr(CO)_5(H_2)$ 364
$Cr(CO)_6$ 351
$Cr(PH_3)_5(H_2)$ 364
Cr–Cr path 367
$[Cr_2(\mu_2\text{-}H)(CO)_{10}]^-$ 367
crambin 288
creation energy 247, 250
critical points 2, 167, 209, 210
$CrOs(CO)_{10}$ M–M bonds 352
CRRK *see* Cash–Karp Runge–Kutta
CRYSTAL 27, 166, 167, 174, 223, 336
– graphs 232
– software 170
– topologies 224
CRYSTAL's LCAO 225
crystalline isostructural families 214 ff.
crystallization 175, 176, 179
crystallographic *R*-factors 334, 337
cubic perovskite $SrTiO_3$ 270
current density 64, 89, 93
curvatures 3, 443
cusp condition 262
cyclic delocalization 404
cycloalkane 99, 114, 116

cyclohexa-1,3-diene 408
cyclohexa-1,4-diene 408
cyclohexane 98, 101, 111, 113, 115, 408
cyclohexene 408
cytosine 415

d
Dalton's atomic hypothesis 39
dangling bonds (DBs) 180
database of protein fragment contents 313
dative bond 350 ff.
DB *see* dangling bonds
DC *see* covalence degree
DCBS *see* dimer-centered basis set
DCD *see* Dewar–Chatt–Duncanson
Debye–Waller factor 286
deformation density 172, 173
– maps 354
delocalization 9, 13–15, 351, 375
π-delocalization 407
delocalization index 13–15, 135, 166, 167, 201, 347, 351, 355, 363, 375, 377, 378, 391, 395, 402, 405, 407, 460
density of states (DOS) 187
depletions 141
Dewar structure 403
Dewar–Chatt–Duncanson (DCD)
– donor–acceptor complexes 357, 361, 363
– ring complex 359
– mechanism 351
– model 364
DI *see* delocalization index
diagonalized quadrupole tensor 24
diamagnetic spin–orbit interactions 378
diamond 208, 216, 219
dielectric polarization 20
Diels–Alder reaction 409, 418 f.
1,12-difluoro[4]helicenes 5, 6
diffuseness *D* 183
dihydrogen bond 426, 456
dihydrogen bonding 8
dihydrogen complexes 364
dimer-centered basis set (DCBS) 460
dimethylcyclobutane 115
dimethyl-phosphinoyl (methylsulfonyl) methane 390
dipeptide L-phenylalanyl–L-proline H_2O 318
dipole derivative 80
dipole moment 21, 22
– derivatives of CO_2 81
dipole polarizability tensor 74
Dirac–Slater exchange density 272
diradicals 393
dispersion energy 459
dissociation energy 436, 439, 440
disynaptic basins 145, 438
DL-serine 329, 334
DMACB 170, 171
DMSDA *see* mean-square displacement amplitudes
DNA bases 25
docking applications 314
domain-averaged Fermi hole 487
σ-donation 352
donor–acceptor bond 350
donor–acceptor interaction 352
DOS *see* density of states
Drude model 215

e
ECP *see* effective core potential
EF *see* eigenvector following
effective core potential (ECP) 27
Ehrenfest
– force 53, 64, 260, 381
– theorem 52
eigenvector following (EF) 168
electric dipole derivative 80
electric dipole moment 68, 71
electric dipole polarization 61
electric field 78, 109, 110
– derivative 73, 75
– flux 63
electric polarizability 68, 73 ff., 90
– tensor 76
electric susceptibility 50
electron delocalization 13–15, 376, 382, 385, 395, 399, 400–401
electron isodensity maps 391
electron localization 9, 14, 128, 143, 401, 440
electron localization function (ELF) 142 ff., 191, 277, 438
electron pair localization 347
electron transfer 431
electron tunneling microscope 54
electronegativity 20, 150, 215, 294, 218, 313, 311, 484
electronic current density 90
electronic energy of an atom in a molecule 82
electronic pressure density 83
electronic structure of molecules 54
electrophile attack 249
electrostatic force 86
– on a nucleus 87
electrostatic interaction energies 312, 313

electrostatic model 351
electrostatic moments 20–25, 138, 139
electrostatic potential 249, 311, 400, 475, 482
electrostatic properties 287, 289, 305 ff.
– of a protein site 313
electrostatic–covalent hydrogen bond model 459
ELF *see* electron localization function
– ELF basins 149 ff.
– ELF population analysis 147 ff.
– ELF topology 144 ff.
– ELF valence basin 166, 200
ellipticity 12
elpasolites 214
energy density 12 ff., 272
energy derivatives 84
energy-gradient-based force 64, 65, 86
– on the nucleus 83
enzymes 511
estrone hormone 15
ethane 98, 109, 112, 113, 194, 379, 380, 418
ethanol 332
ethene 98, 111
ethyne 98, 111
Euler equation 41
Euler's invariant formula 211
exact exchange energies 133
exact exchange force 127
EXAFS *see* X-ray absorption fine structure
exchange between electrons of the same spin 272
exchange correlation 377
exchange density 274
exchange eigenvalues 138
exchange energy 122, 124, 125, 127, 128, 130 ff., 136, 276, 277
exchange force 136, 137
exchange interaction 131
exchange interaction energy 461
exchange moments 126, 139
exchange potential 278
exchange-correlation density 401, 402
exchange-correlation energy 271
expectation value of an operator 9
experimental electron density 261
experimental H···O interactions 436
external magnetic field 89
external potential 474
EXTREME 27

f

F_2 445
F-center in sodium electrosodalite 190
F-center basins 192
F-centers 186
Fe–Fe bonding 369
Fermi
– contact 378
– contact contribution 379
– – to $^3J_{HH}$ 378
– correlation 14
– hole 144, 167, 375, 377, 378, 402, 403, 487
– level 189
ferromagnetic phase of SES 191
Feynman path–integral method 271
Feynman
– force 52, 381
– theorem 52
F–F coupling constants 380
F···H hydrogen bonded complexes 441
$[F\cdots H\cdots F]^-$ system 444
F–H···π hydrogen bond path 456
FH···ClH 146
$(FH)_2$ 420, 437
first atomic electrostatic moment 20
first derivative of the electric dipole moment 78
first-order reduced density matrix 123
Fisher information 144
fixed nucleus approximation 55
flatness 354
π-fluctuation aromaticity index (FLU_π) 406 ff., 415, 419, 420
flux through a surface 47
force fields 26, 121
formamide 333
formic acid–formate anion complex 199
Fourier
– difference synthesis 225
– transformation 289
fpLAPW *see* full potential linearized plane wave
FRAGDIP 21
fragment charges 312
fragment deformation maps 354
fragment electroneutrality 307
fragment representations 313
fragment transferability 338
free electron gas 215
Fukui
– frontier orbital theory 54
– functions 475
– Nobel Lecture XXI
– radical reactivity indices 484
full potential linearized augmented plane wave (fpLAPW) formalism 223

fullerenes 319, 409
Fulton bound index 391

g
GAMESS 27, 29, 128
gauche effect 383
Gauss's theorem 18
Gaussian 27, 29, 30, 106, 114 ff., 332
Gaussian 03 129
Gaussian 94 295
genetic code 22–24, 26
GFMLX 336
GIAO 413
globbic structure factors 289
glycine 124
goodness of fit 334
gradient kinetic energy 17
gradient path 242, 428, 474
gradient vector 6
– field 6, 142, 259, 323
– field lines 211
– instability 210
gradient-corrected correlation energy 275
gradient-corrected energy density 272
gradient-corrected exchange density 274
graphite 208, 219
GRDVEC 28, 29
Green's function 192, 263, 271
GRIDV 28, 29
group additivity schemes 82
group contribution 22, 50
– to the polarizability tensors 77
guanine 415, 416
guanine–cytosine base pair 415
guest atoms 188
guest–host binding energy 189
guest–host systems 165 ff.
– binding energy 189

h
$H_2N–C_\alpha H_\alpha(R)–COOH$ 21–23, 295
$H_3N\cdots HF$ 439
H_2O dimer 129
Hamiltonian approach to quantum mechanics 38
Hammett substituent constants 412
Hansen–Coppens XXII
– multipole model 261, 262, 318
hAR *see* human aldose reductase
harmonic approximation 78
harmonic oscillator model of aromaticity (HOMA) 394, 400, 404, 406 ff., 410, 412 ff., 420
Hartree–Fock
– energy 123
– virial 29–30
– wave functions 127
Hattig's recurrence formulae 129
HB *see* hydrogen bond
HCCH···HF complex 456
HCN 445
HCH···O intermolecular interactions 170
HDN *see* hydrodenitrogenation
HDS *see* hydrodesulfurization
heats of formation 50
heavy main-group element 352
hedrane 114, 116
Heisenberg
– equation of motion 37
– representation of quantum mechanics 51
Hellman–Feyman
– electrostatic force 19, 65, 83
– electrostatic theorem 83, 85, 87, 88
Hermitian operator, linear 46
Hermiticity 47
Hessian matrix 3, 231, 427
hexaprismane 117
1,3,5-hexatriene 129, 134
HF dimer 129, 130
H···F hydrogen-bonding interactions 437
H···H bonding 8, 9, 11
H···H electrostatic interaction 445
hierarchical merging/clustering algorithm 292, 301
higher-order polarizability 104
high-pressure phosphorous boride 215
high-resolution
– electrostatic potentials 306, 308, 309
– protein model 287
Hirshfeld
– multipole model 262
– test 318, 335
H···N hydrogen-bonding interactions 441
$(H_2O)_2$ 131, 132, 430
H···O hydrogen bonds 436
H···O interactions 437
Hohenberg–Kohn
– formulation of DFT 261
– theorem XX, 474, 475
HOMA *see* harmonic oscillator model of aromaticity
homodesmotic reaction 400
homogeneous electron gas 263
homoleptic $M_2(CO)_n$ dimers 352
host–guest
– chemistry 186
– systems 186 ff.

human aldose reductase (hAR) 287 ff., 291 ff., 300
– crystal structure 299
– structure 307, 312, 313
hybrid orbital-free energy functionals 271
hydride bond 466
hydride bridges 367 ff.
hydrides 153
hydrodenitrogenation (HDN) 236
hydrodesulfurization (HDS) 236, 237
hydrogen bond(ing) (HB) 170, 177, 197, 199, 217, 273, 331, 416, 425, 427, 429, 453
– Coulomb interaction 130
– donor 504
– energies 454
– exchange energy 132
– intermolecular 453
– molecular complexes 198
hydrogen–hydrogen bonding 8
hydroimidazo[4,5-*d*]imidazole 430
hydrophilic regions 500
hydrophobic effects 481
hydrophobic regions 500
hydrophobic–hydrophilic interaction tendency 483
hydrophobicity 483
hydroxy-2-aminopurine 508 (6-)
hyperconjugation 386
hyperpolarizability 61, 73
hypervirial theorem 83, 86
hyperwall mode 505

i

I_2O 216
IAM *see* independent atom model
ice VIII 436
iceane *see* tetracyclo-[5.3.1.1^{2,6}.0^{4,9}]-dodecane
independent atom model (IAM) 172, 174, 189, 286 f., 329, 334, 336
independent transferability 113
induced electronic magnetic dipole moments 88 f.
influence function 192
infrared intensity 50
infrared spectrum 78
infrared vibrational absorption intensity 68
inhibitor–protein interactions 286
inner-valence shell charge concentration (i-VSCC) 509 ff.
inorganic clathrates 186
integrated atomic electronic energy 277
integration error 18, 276
intensity of absorption 78
interaction
– density 261
– energy 435
– energy-decomposition scheme 459
– potential 436
– tensor 127
interatomic surface (IAS) 8, 11, 63, 244, 232, 377, 426, 477
intermolecular interaction 131, 435
invariom 317, 331, 334
– aspherical scattering factors 329
– database 330, 336
– pseudoatoms density 338
inverse hydrogen bonding 466
inverse moments 122
ionic bonds 217
ionic character 151
ionic contributions 151
ionicity 216
isobutane 99
isodensity envelope 17
isopropanol 332
isotropic polarizability 77

k

Karplus-type behavior 379, 382
– of $^3J_{HH}$ 395
Kekulé
– resonance structure 416
– structures 403
Kenichi Fukui XXI
kinetic energy 43 f., 122, 123, 266
– density 12, 13, 18, 37, 259, 262 ff., 349, 354, 381, 391, 432, 482, 483
– operator 17
– per electron 192
Kohn–Sham
– approximation 378
– density-functional theory 128
– exchange-correlation energy 128
– orbitals 128

l

Lagrange polynominal interpolation 232
Lagrangian
– action principle 37
– classical 41
– equation of motion 41
– representation of quantum mechanics 51
L-alanine 208
Laplacian of the electron density 3, 10, 17, 44, 141, 149 f., 177, 182, 192, 200, 215, 218,

Laplacian of the electron density (cont.) 259, 262, 324, 346, 363, 427, 445, 458, 483, 488, 500 ff.
LBHB *see* low-barrier hydrogen bonds
Le Chatelier principle 56
Leu–Enkephalin 294
Lewis, G. N. 502
– acid 10, 252, 253, 466
– acidity VSCCs 367
– base 10, 466
– model 145, 345, 352, 378, 438
– structures 157
LFT *see* ligand-field theory
Li 216
LiC≡CLi···HF complex 456
LI-DICALC 28
LIF 269, 270
ligand–protein interactions 286
ligand-field theory (LFT) 349, 352
limits of transferability 331
local aromaticity criteria 413
local bulk moduli 220, 221
local dipole moment 96
local electron affinity 484
local electronic charge concentration 10
local energy densities 443
local exchange energy 273
local hardness 484
local polarizability 484
local source (LS) 192, 194
local statement of the virial theorem XX, 13, 56, 346
localization
– domains 145 ff., 156
– function 142
– indices 14, 15, 167, 377
– nodes 146
– tree-diagram of H_2CO 147
– tree-diagram of NaCl 147
lock and key mechanism 352
lone pair 145, 438
– electron concentrations 275
Lorentzian form 144
low-barrier hydrogen bonds (LBHB) 446
low-electronegativity elements 214
low-pressure boron phosphide 215
LS *see* local source

m

M···CO interactions 352
$M_2(CO)_n$ 356
macromolecular crystallography 285
magnetic dipole moment 88
magnetic susceptibility 50
magnetizability 64, 68, 90
– tensor of naphthalene 92
malonaldehyde 199
matrix of exchange moments 136 ff.
maximum entropy method (MEM) 187, 188
M–C interactions 370
mean molecular polarizability 97, 98
mean-path approximation 271
mean-square displacement amplitudes (DMSDA) 335
measurable properties 50
medium-resolution electrostatic potential 310
MEM *see* maximum entropy method
metallacycles 357
metal–metal (M-M) bond(ing) 199, 349, 352, 354 ff., 370
– bond path 354, 369
– contacts 352
– interactions 354, 370
metal–olefin complexes 361
methane 107, 109, 113
methanol 332
methyl fluoride 509
methyl group 77, 195
N-methylacetamide 333
methylcycloalkanes 99
methylcyclobutane 115
N-methylformamide 333
methylcyclopropane 115
methylene group 56, 109
misdirected valence 511
Mn···(CO) intramolecular interactions 356
$Mn_2(CO)_{10}$ 201, 355
Mn–Mn bond 200, 201, 355
molar volumes 50
molecular complexes 196 ff.
molecular crystals 170 ff.
molecular devices 55
molecular dipole 170, 176
molecular electrostatic potentials 313
molecular expectation value 10
molecular graph 8, 52
molecular orbital theory 54
molecular polarizability 73, 77, 96ff., 99, 109, 110, 113
molecular potential 122
molecular quantum similarity measures (MQSM) 487
molecular response properties 61
molecular similarity 26
molecular structure stability 52
molecular virial theorem XX, 29, 86, 88
molecular volume 481

molecular graph 5
Møller–Plesset perturbation theory 29
monopolar exchange moment 128
monosynaptic basins 145, 438
MORPHY 27, 332
– 01 128, 129
Morse equation (relationship) 4, 168, 172
Morse-type potentials 437
Mo–S
– bond 242, 244, 246
– inter-atomic surfaces 246
MoS_2 239, 245 ff.
– bulk 241 ff., 244
– unit cell 241
MQSM *see* molecular quantum similarity measures
mRNA codon 22–24
muffin tins 223
multiple bonds 156 ff.
multiple-exchange energies 133
multipolar database library 287, 306, 307, 309
multipolar ED database fragment 287
multipolar refinement technique XXII
multipole expansion 130 ff., 137
multipole model 264, 277, 336
multipole moments 121

n

N···N contacts 172
N_2 216
N_2···HF 439
NaCl 146
$NADP^+$ 293, 302, 312
NaF 348
nanotechnology 55
naphthalene 91, 93, 392 ff., 408
naphthacene 408
natural bond orbital (NBO) 446
natural coordinates 104
natural orbitals 378
NBM *see* non-bonded maxima
NBO *see* natural bond orbital
Ne_2 348
nearest-neighbor approximation (NNA) 333
negative hyperconjugation 364
net current vector 64
neural networks 121
Newton's equation of motion 41
Newton–Raphson (NR) technique 168, 231, 236
N···H hydrogen-bonding 433
N–H···N hydrogen bonds 172, 446
– complexes 441
N–H···O hydrogen bonds 172, 448, 462
NICS *see* nucleus-independent chemical shift
NiMoS 239, 240
Ni–Ni interaction 354
NMR shielding tensors 90
N–N
– bond 217
– contacts 171, 172
NNA *see* non-nuclear attractors *and* nearest-neighbor approximation
non-bonded charge concentration (NBCC) 182
non-bonded interactions 130
non-bonded maxima (NBM) 177, 178
non-nuclear attractors (NNA) 6, 168, 191–193, 208, 314
non-nuclear maxima (NNM) 6, 216, 218, 314, 225, 226
non-stationary point geometry 62, 83, 84
normal mode vibration 80
normalized spherical harmonics 126
normal-mode vibrational coordinates 79, 80
NR *see* Newton–Raphson
N-representability problem 378
nuclear critical points 2, 4, 168
nuclear momentum operator 86
nuclear virial energy 62, 68, 83, 86
nucleophilic attack 249, 250
nucleus-independent chemical shift (NICS) 400–420
null (zero value) molecular property 61, 62, 64, 66, 70
numerical integration error 29

o

observables 16
occupation numbers 378
octet rule 155
OH fragment transferability 332
O–H···O hydrogen bond 448, 462
oligopeptide molecules 336 ff.
one-electron density matrix 56, 123, 166, 260, 263
open quantum system 44
open-shell character 355
orbital conservation 54
orbital models 54
ORCRIT 289, 290
origin-dependent atomic charge-transfer dipole contribution 72
origin-dependent atomic contributions 61, 68, 87
origin-dependent atomic property 85

origin-dependent charge transfer term 62
origin-dependent polarization term 62
origin-independent atomic contribution 61
origin-independent property density 61
overlap function at a point 125

p

packing forces 368
PAH *see* polycyclic hydrocarbon
pair density 166, 376
– functions 376
para-delocalization index (PDI) 404 ff., 410, 414, 415, 417
paramagnetic spin–orbit 378
para-nitroaniline 71 ff., 74, 76 ff.
PASA *see* promolecular atom shell approximation
path integral approach 49
Pauli
– exclusion principle 377
– repulsion 144, 381, 438
PBH *see* polybenzenoid hydrocarbons
PCl_5 155
PDI *see* para-delocalization index
penamecillin 499, 504
penicillin 506, 507
– derivative 504
n-pentacosane 114
pentadecane 114
pentane 111, 113, 115, 194
peptide
– bond 328
– crystals 297
– HN–$H_\alpha C_\alpha$–C=O group 299
– plane HN–$H_\alpha C_\alpha$–C=O 294
pericyclic concerted reaction 418
periodic systems 231
perovskites 214
pharmacophores 499
phenanthrene 8, 392 ff., 408, 416
phenol 332
pK_a of weak acid 26
Poincaré–Hopf formula (relationship) 4, 6 f., 66, 209
point-charge models 122
polar bonds 215, 455
polarizability 64, 78, 80, 110, 111, 400, 483
– tensor 11103
polybenzenoid hydrocarbons (PBHs) 391, 393, 394
– aromatic dilution 394
polycyclic aromatic hydrocarbons (PAHs) 404, 409, 410, 416, 417
polymorphism 208
polysynaptic basins 438
population analysis 145
porphyrin 288
post Hartree–Fock wave functions 27–30, 127
potential energy density 8, 12, 25, 56, 259, 262 ff., 349, 354, 381, 462
powerwall mode 505
primary bundle 211, 212
principle
– of least action 38
– of stationary action 47, 49
PROAIM 27, 29, 193, 477
procrystal 267
PROMEGA 193
promolecular atom shell approximation (PASA) 290
promolecular model 217, 222
promolecule maps 300
propane 98, 109 ff., 194
propellane 99, 114
proper open quantum system 7, 8, 50
proper operator 50
properties after molecular reconstruction 481
protein
– binding site 313
– crystal structures 285
– crystallography 285
– electrostatic properties 287
– main chain HN–C=O peptide plane 288
– molecules 336 ff.
– refinement 289
– retention time 26
– stability 26
– structure 289, 311
protein–DNA docking 290
protein–ligand interaction energies 289
protein–protein docking 290
protein–protein interaction energies 289
proton–proton vicinal coupling constants 380
pseudoatom 7, 275, 329
– fragments 317
pseudoatomic density 261
pseudopotential 169, 222
pyracelene 408
pyridine 406, 408
pyrimidine 408

q

QTAMC *see* quantum theory of atoms in molecules and crystals

quadrupole moment 24
quality of an atomic integration 18
quantitative structure–activity/property relationship 473
quantum
– observables 37
– self-similarity measure 487
– stress tensor density 51
– vector current density 51
quantum theory of atoms in molecules and crystals (QTAMC) 259, 272
quinoline 408

r
radial density functions 261
radius concept 207
Raman
– scattering intensity 95, 97 ff., 103 ff.
– spectroscopy 95
rank 4
reciprocal space vectors 289
RECON 476 ff.
reduced density matrices 121
reference density 188
regularity in the genetic code 24
Reiss–Münch theorem 475
relative hardness 400
relativistic effects 55
relief map of the electron density 2
residual electron density 318, 337
resonance structures 416
resonance-assisted hydrogen bonds 457, 458
response properties 62
R-factor 336
ring
– critical point (CP) 4, 5, 168, 388
– currents 399
– strain 5
– surface 5
rotational barriers 375

s
scanning tunneling microscopy (STM) 237
Schrock carbene 512
Schrödinger's
– functional 46
– kinetic energy 17
– time dependent equation 41
Schwinger's principle of stationary action 8, 46, 50, 476
scorpion toxin 288
SCVS *see* self-consistent virial scaling
SD *see* softening degree
second atomic electrostatic moment 24
second-order (pair) density 401
second-order density matrix 14
second-order Jahn–Teller symmetry rule 54
self-consistent virial scaling (SCVS) 30, 85, 86
self-interaction 124
separability 394 (σ–π)
SES *see* sodium electrosodalite
SF_4 155
SF_6 155
shared-shell
– character 445
– interactions 174, 441, 443
sharing of electrons 13
shell structure 346
SHELXL 336
short strong hydrogen bond (SSHB) 463
Si crystal 225, 226
Si(111)(1 × 1) 180 ff., 184
Si(111)(1 × 1)–H 181
– surface 180
Si(111)(2 × 1) 180
– reconstructed surface 184 ff.
– surface 185
Si_2Me_2 157
Si_2Me_4 157
side-chain fragments 308
signature 4
silanes 114, 117 f.
silicon bulk 182
similarity index 26, 488
single-enantiomer drugs 336
singlet diradicals 393
Si–O bonds in silicates 436
Si–Si bond properties 180
slab model 180, 181
sodium electrosodalite (SES) 186, 190 ff.
softening degree (SD) 447
solid peoperties 207
source contribution 198, 201
source function 165, 192 ff., 196 ff., 201 f.
spherical harmonics 125
spin density 148, 190, 191
spin population 16, 143
spin–dipolar interactions 378
spin-independent electron correlation 272
spin-less pair density 401
S–S
– bond 242, 251, 427
– bond critical points 242, 245
– bond paths 242
– interatomic surface 253
SSHB *see* short strong hydrogen bond
stability of gradient vector field 53

π–π stacking interaction 25
STM *see* scanning tunneling microscopy
stress tensor 64
strong hydrogen bonds 455
strong van der Waal's covalent interactions 440
structural stability 52, 53
substituent effects 412
supramolecular chemistry 186
surface
– contribution 46
– derivatives 75
– flux 47
– layer 182
– terms 46
symmetry-stabilized agostic interaction 368
synaptic order 145, 166, 438
synchrotron radiation 318 ff.

t

TAE *see* transferable atom equivalent
tert-butylcyclohexane 391
tetalin 392
tetracyclene 99
tetracyclo-[5.3.1.1^{2,6}.0^{4,9}]-dodecane (iceane) 115
tetrahydroimidazol[4,5-*d*]imidazole 429
tetralin 395
thermodynamic property 208
– of ice VIII 437
– partitioning 220
Thomas–Fermi
– approximation 263
– kinetic energy functional 438
three-center (*3c*) systems 357
three-center bonding 356 ff.
$TiCl_3(C_2H_5)$ 366
time-independent electric field 74
TMS *see* transition metal sulfides
topological atoms 123, 129
topological multipole expansion 122
topological polymorphism 215
topological polytypism 208, 214, 215
topological properties 376
TOPOND 27, 165 ff., 170
TOPXD 168 ff.
total energy 123
total energy density 346, 432
total molecular volume 178
total polarization contribution 73
transferability 55, 56, 82, 111, 194, 287, 294, 297 f., 317, 324, 327, 331, 332, 504
transferable atom equivalent (TAE) 474 ff., 513
transferable methylene contribution 77
transferable multipolar data 289
transition metal sulfides (TMS) 236
transition probability 26
transition metal
– atoms 509 ff.
– carbonyl complexes 351
– ions 511
triazine 408
triphenylene 408
trisynaptic basins 145, 438
TS *see* aromatic transition state
two-center bonding 349 ff.
two-center–two-electron (*2c–2e*) bonds 349
two-electron density matrix 166, 260
tyrosine 22, 24, 294, 299
tyrosine–glycine peptide bond 294

u

uniform electron gas 272
unsaturated hydrocarbons 117 f.
urea 171 ff., 264 ff.
– crystal 171, 177

v

valence basin 142, 438
– populations 149
valence flatness 315
valence shell charge concentration (VSCC) 149, 177, 182, 347, 349, 350, 355, 502
valence shell charge depletion (VSCD) 502, 508
valence shell electron pair repulsion (VSEPR) model 141, 149, 150, 153 ff., 157, 158, 378, 511
– electron domains 149 ff.
van der Waals
– bonds in graphite 219
– complexes 454
– interactions 455
– repulsion 383
variation of the surface 45
vibrational frequency 84, 96
vibrational modes 96
vibrational spectra 79
virial 30
virial field 12, 56
virial graph 8
virial of the electronic Ehrenfest force 83
virial operator 83
virial path 8, 9, 260
virial ratio 20, 29
virial theorem XX, 20, 29, 37, 52, 85, 260, 271, 432

virtual high-throughput screening 473
vitamin B_{12} 320, 321, 323
VmoPro 307, 309, 311
volume rendering 500
volumes 150, 178
von Weizäcker kinetic energy functionals 438
VSCC *see* valence shell charge concentrations
VSCD *see* valence shell charge depletion
VSEPR *see* valence shell electron pair repulsion

w
$W(CO)_3(PR_3)_2(H_2)$ 364
Walter Kohn XXII
water 276
water dimer 137, 138, 199
Watson–Crick base pair 415
wavefunction 27, 42, 85
weak hydrogen bonds 455
weak interactions 441
weak van der Waal's covalent interactions 440
WIEN 223
WIEN's fpLAPW 225

x
XD 165, 168, 169
X–H···F–Y complexes 443, 444
X–H···O hydrogen-bonding interactions in crystals 436
X–H···Y hydrogen bonds 499
$(XH_3)(CO)_3Co–Co(CO)_3(XH_3)$ molecules 355
X-ray absorption fine structure (EXAFS) 236
XTAL 289, 290, 298, 300, 303

z
zero flux
– condition 6, 9, 17, 44, 46, 275
– surface 6, 9, 10, 44, 48, 323, 354, 426, 429

Related Titles

Rode, B. M., Hoder T., Kugler, M.

The Basics of Theoretical and Computational Chemistry

2007
ISBN-13: 978-3-527-31773-8
ISBN-10: 3-527-31773-2

Mandl, F., Shaw, G.

Quantum Field Theory

2007
ISBN-13: 978-0-471-49683-0
ISBN-10: 0-471-49683-9

Bethge, K., Gruber, G., Stöhlker, T.

Physik der Atome und Moleküle
Eine Einführung

2004
ISBN-13: 978-3-527-40463-6
ISBN-10: 3-527-40463-5

Demtröder, W.

Molecular Physics
Theoretical Principles and Experimental Methods

2004
ISBN-13: 978-3-527-40566-4
ISBN-10: 3-527-40566-6

Kaupp M., Bühl M., Malkin V. G. (Eds).

Calculation of NMR and EPR Parameters
Theory and Applications

2004
ISBN-13: 978-3-527-30779-1
ISBN-10: 3-527-30779-6